National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-59587-0

Canada

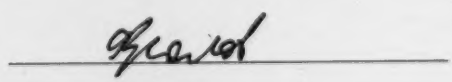
University of Alberta

Library Release Form

NAME OF AUTHOR:	Andrei V. Frolov
TITLE OF THESIS:	Critical Behavior in Gravitational Collapse of the Scalar Field
DEGREE:	Doctor of Philosophy
YEAR THIS DEGREE GRANTED:	2000

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Andrei V. Frolov
Physics Department
University of Alberta
Edmonton AB Canada, T6G 2J1

6 September 2000

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Critical Behavior in Gravitational Collapse of the Scalar Field** submitted by **Andrei V. Frolov** in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Don N Page

Prof. D.N. Page (Supervisor)

Prof. M. Razavy (Chair)

Prof. M.W. Choptuik (External)

Prof. D.P. Hube

Prof. H. Kunzle

Richard Sydora

Prof. R. Sydora

DATE:

Abstract

Critical phenomena in gravitational collapse have been a relatively recent and interesting development in the established field of general relativity. Due to the difficulty of the problem, most of the work on the subject is done using numerical methods. In my study, I approach critical collapse of a massless scalar field using analytical methods. I consider evolution of scalar field configurations close to a certain continuously self-similar solution at the threshold of black hole formation, and analyze their departure from that solution at the later times using perturbative methods. In the framework of this analytical model, one can observe (and test) important general features of the critical collapse, such as universality of the critical solution with respect to initial data, mass scaling in black hole production, and formation of discretely self-similar structure in the course of the field evolution.

Contents

1 Introduction 1

2 Overview of Critical Phenomena in General Relativity 4
- 2.1 Critical collapse of a massless scalar field . . . 4
- 2.2 Generic features of critical collapse . . . 6
- 2.3 Self-similarity . . . 9
- 2.4 Linear perturbation analysis and mass scaling . . . 11
- 2.5 Literature overview . . . 13

3 The Generalized Roberts Solution 16
- 3.1 Spherically-reduced action and field equations . . . 17
- 3.2 n-dimensional self-similar solutions . . . 18
- 3.3 Critical behavior . . . 21
- 3.4 Particular cases . . . 22
  - 3.4.1 $n = 3$ . . . 22
  - 3.4.2 $n = 4$ . . . 23
  - 3.4.3 $n = 5, 6$ . . . 24
  - 3.4.4 Higher dimensions . . . 24
- 3.5 General scalar field coupling . . . 25
  - 3.5.1 Equivalence of couplings . . . 25
  - 3.5.2 Examples . . . 27
- 3.6 Properties of the Roberts solution . . . 28
  - 3.6.1 Global structure . . . 29
  - 3.6.2 Newman-Penrose formalism . . . 30
  - 3.6.3 Scaling coordinates . . . 31
  - 3.6.4 Curvature coordinates . . . 32
- 3.7 Switching the scalar field influx on and off . . . 33
  - 3.7.1 Matching with flat spacetime . . . 33
  - 3.7.2 Matching with outgoing Vaidya solution . . . 34
- 3.8 Quantum corrections . . . 38

4 Spherically Symmetric Perturbations **40**
4.1 Gauge-invariant linear perturbations 40
4.2 Decoupling of perturbation equations 42
4.3 Separation of variables 44
4.4 Growing modes 45

5 Beyond Spherical Symmetry **49**
5.1 Gauge-invariant linear perturbations 49
5.2 Perturbation equations 51
5.3 Decoupling of perturbation equations 54
5.4 Perturbation spectrum 56

6 Continuous Self-Similarity Breaking **58**
6.1 Wave propagation on the Roberts background 58
6.1.1 Outgoing wavepacket 60
6.1.2 "Constant" Wavepacket 61
6.1.3 Incoming Wavepacket 61
6.2 Late-time behavior of incoming wavepacket 64
6.2.1 Evolution near $v = 0$ 66
6.2.2 Evolution of a wavepacket initially localized at $v = 0$ . . . 67
6.2.3 Generic initial conditions 69
6.3 Emergence of discrete self-similarity 70

7 Discussion **74**
7.1 Our results 74
7.2 Further studies 75

Appendices **77**

A Properties of the Hypergeometric Equation **78**

Bibliography **80**

Chapter 1

Introduction

Critical phenomena in gravitational collapse have been a relatively recent and interesting development in the established field of general relativity. Following the numerical work of Choptuik on the spherically symmetric collapse of the minimally coupled massless scalar field [31], critical behavior was discovered in most common matter models encountered in general relativity, including pure gravity [1], null fluid [37] and, more generally, perfect fluid [81, 88], as well as more exotic models.

The essence of critical phenomena in general relativity is the fact that just at the threshold of black hole formation, the dynamics of the field evolution becomes relatively simple and, in some important aspects, universal, despite the complicated and highly non-linear form of the equations of motion. In analogy with second order transitions in condensed matter physics, the mass of the black hole produced in near-critical gravitational collapse usually scales as a power law,

$$M_{\rm BH}(p) \propto |p - p^*|^\beta, \tag{1.1}$$

with parameter p describing initial data. The mass-scaling exponent β is dependent only on the matter model, but not on the initial data family. The critical solution, separating solutions with black hole formed in the collapse from the ones without a black hole, also depends on the matter model only, and serves as an intermediate attractor in the phase space of solutions. It often has an additional symmetry called self-similarity, in either continuous or discrete flavors.

The discovery of critical phenomena in gravitational collapse was a real success of numerical relativity in which a physical effect was observed in simulations without being first predicted by theoretical physicists. For the theoretician, the challenge and attraction of studying critical phenomena lie in the possibility of exploring a new class of exact solutions of Einstein's equations, having simple properties and high symmetry, but previously undiscussed. Another interesting thing about critical solutions is that they are relevant to the cosmic censorship

conjecture, offering a rather generic counterexample of a naked singularity forming in gravitational collapse of regular data.

Much work has been done, and most general features of critical behavior are now understood. However, the theory of critical phenomena in general relativity is somewhat incomplete with respect to our lack of knowledge of critical solutions (or even their approximations) in explicit closed form. Due to the obvious difficulties in obtaining solutions of Einstein equations, most of the work on critical phenomena seems to be using numerical methods. In this thesis, we attempt to remedy this deficiency by studying critical collapse of a massless scalar field using analytical models and methods.

Layout of the thesis

This thesis is organized in the following way:

Chapter 2 is an introduction to critical phenomena in general relativity. It describes generic features of critical behavior in general relativity, presents evidence in support of universality of critical collapse, and discusses self-similarity and its role in critical solutions. A review of the extensive literature on the subject is provided in Section 2.5.

Chapter 3 constructs analytical solutions of a spherically symmetric, continuously self-similar gravitational collapse of a scalar field in n dimensions. The qualitative properties of these solutions are explained, and the critical behavior in this model is discussed. Closed-form answers are provided where possible. Equivalence of scalar field couplings is used to generalize minimally coupled scalar field solutions to the model with general coupling.

Chapter 4 considers stability of the continuously self-similar critical solution of the gravitational collapse of a massless scalar field which lies at the threshold of black hole formation (the Roberts solution). The linear perturbation equations are derived and solved exactly. The perturbation spectrum is found to be not discrete, but occupying a continuous region of the complex plane.

Chapter 5 continues the perturbative analysis of Chapter 4 beyond spherical symmetry. The exact analysis of the perturbation equations reveals that there are no growing non-spherical perturbation modes. This shows that all the non-sphericity of the initial data decays in the collapse of the scalar field, and only the spherically symmetric part will play a role in the critical behavior.

Chapter 6 considers the near-critical evolution of the spherically symmetric scalar field configurations close to the continuously self-similar solution in more detail. Using Green's function method, it is shown that a generic growing perturbation departs from the Roberts solution in a universal way. We argue that in the course of its evolution, initial continuous self-similarity of the background is bro-

ken into discrete self-similarity with echoing period $\Delta = \sqrt{2}\pi = 4.44$, reproducing the symmetries of the critical Choptuik solution.

Chapter 7 concludes this thesis by discussing the obtained results. We briefly summarize the material presented here and ponder the possibilities for future studies.

Conventions

Throughout this thesis we use units where $G = c = 1$. Greek letter indices run through $0, 1, 2, 3$, capital Latin indices take values $\{0, 1\}$, and lower-case Latin indices run over angular coordinates $\{2, 3\}$. A semicolon (;) denotes the covariant derivative in four-dimensional spacetime, and a stroke (|) denotes the covariant derivative in two-dimensional spacetime with respect to the two-metric γ_{AB}. The sign conventions are as in Misner, Thorne, and Wheeler's book *Gravitation*. The Newman-Penrose formalism quantities follow the original sign conventions of Newman and Penrose.

Chapter 2

Overview of Critical Phenomena in General Relativity

In general, there are two distinct possible late-time outcomes of gravitational collapse, classified by whether the black hole was formed in the collapse or not. The first end state consists of a black hole, plus some outgoing matter and outgoing gravitational radiation. The second end state consists possibly of a stationary remnant star (if it is allowed by the matter model), plus some outgoing matter and outgoing gravitational radiation, but no black hole. Which end state is realized for a particular evolution depends on initial conditions: gravitational collapse produces a black hole only if the gravitational field becomes strong enough during the collapse. At the threshold of black hole formation critical phenomena occur.

In this chapter, we talk about generic features of critical behavior in general relativity, and review the extensive literature on the subject.

2.1 Critical collapse of a massless scalar field

In numerical simulations of the spherically symmetric gravitational collapse of a minimally coupled massless scalar field, Choptuik discovered a new and exciting physical effect — the existence of critical phenomena in general relativity [31]. In this section, we briefly summarize his findings.

The general time-dependent, spherically symmetric metric can be written in Schwarzschild coordinates as

$$ds^2 = -\alpha^2(r,t)\,dt^2 + a^2(r,t)\,dr^2 + r^2\,d\Omega^2, \tag{2.1}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta\,d\varphi^2$ is the metric of a unit two-sphere. The evolution of

the minimally coupled scalar field ϕ is described by Einstein equations[1] and the massless Klein-Gordon equation,

$$R_{\mu\nu} = 8\pi\phi_{,\mu}\phi_{,\nu}, \quad \Box\phi = 0, \tag{2.2}$$

which for metric (2.1) have the form

$$\frac{\partial}{\partial t}\left(\frac{a}{r}X\right) = \frac{\partial}{\partial r}\left(\frac{\alpha}{r}Y\right), \tag{2.3a}$$

$$\frac{\partial}{\partial t}\left(\frac{a}{r}Y\right) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r\alpha X\right), \tag{2.3b}$$

$$\frac{1}{\alpha}\frac{\partial\alpha}{\partial r} - \frac{1}{a}\frac{\partial a}{\partial r} + \frac{1-a^2}{r} = 0, \tag{2.3c}$$

$$\frac{1}{a}\frac{\partial a}{\partial r} + \frac{a^2-1}{2r} - \frac{a^2}{r}(X^2+Y^2) = 0, \tag{2.3d}$$

where auxiliary scalar field variables X and Y defined by

$$X(r,t) = \sqrt{2\pi}\,\frac{r}{a}\frac{\partial\phi}{\partial r}, \quad Y(r,t) = \sqrt{2\pi}\,\frac{r}{\alpha}\frac{\partial\phi}{\partial t} \tag{2.4}$$

were introduced. Both system of equations (2.3) and auxiliary variables (2.4) are invariant under simultaneous rescaling of time and space variables $r \mapsto kr, t \mapsto kt$. This invariance reflects the absence of mass/length scale in the model. The solution is most conveniently analyzed in terms of the logarithmic coordinates

$$\rho = \ln r + \text{const}, \quad \tau = \ln(T_0^* - T_0) + \text{const}, \tag{2.5}$$

where T_0 is the proper time of a central observer $T_0 = \int \alpha(0,t)\,dt$, and T_0^* is the moment of formation of the black hole.

Choptuik solved equations (2.3) numerically using a finite-difference technique with adaptive mesh-refinement, where the basic discretization scale h is allowed to vary locally, both in space and time, in response to the development of the solution features. The computations were made for several families of the initial scalar field profiles dependent on the parameter p, and for each family, a critical value p^* corresponding to the threshold of the black hole formation was found using a binary search.

The critical solution was found to be universal for all families and serving as an intermediate attractor for all nearby solutions. Near-critical solutions were

[1] Note the factor of 8π in the Einstein equations used by Choptuik. Elsewhere throughout this thesis, we use a slightly different convention where a factor of 4π is absorbed in the definition of the scalar field action. In our convention, the Einstein equations are written as $R_{\mu\nu} = 2\phi_{,\mu}\phi_{,\nu}$.

observed to display a curious symmetry, known as discrete self-similarity (a.k.a. echoing), where asymptotic critical field profiles (denoted by asterisk) are invariant under rescaling of spacetime

$$\{X, Y\}^*(\rho - \Delta, \tau - \Delta) \simeq \{X, Y\}^*(\rho, \tau) \tag{2.6}$$

by a particular factor $\Delta \approx 3.44$. The mass of the black holes produced in supercritical evolution was remarkably well described by the power law

$$M_{\rm BH}(p) \propto |p - p^*|^\beta \tag{2.7}$$

over a wide range of black hole masses, with exponent being approximately $\beta \approx 0.37$. The numerical results indicated that there is no mass gap in this model, and the black hole production is turned on at infinitesimal mass.

2.2 Generic features of critical collapse

Using $3 + 1$ split of the spacetime, the gravitational collapse is described by the initial value problem, dynamics of which is given by the Einstein equations

$$G_{\mu\nu} = 8\pi T_{\mu\nu}, \tag{2.8}$$

where $T_{\mu\nu}$ is the stress-energy tensor obeying the conservation laws

$$T^{\mu\nu}{}_{;\nu} = 0, \tag{2.9}$$

and equations governing matter dynamics. The initial data for this problem is specified on some past Cauchy surface. As time passes, the solution (given by a matter field and a metric on a three-surface of constant time) is evolving in the configuration space of all solutions. This configuration space is divided into two parts according to the end result of the evolution, with the critical threshold hypersurface separating the "non-black-hole-end-state" solutions from the "black-hole-end-state" ones, as illustrated in Fig. 1. Among the infinite number of parameters describing the initial state, pick a single parameter p characterizing the strength of gravity in the ensuing field evolution.

For small values of parameter p the gravitational field during evolution is too weak to form a black hole, while for large values of p the black hole is produced. In general, between these two extremes there is a *critical* parameter value, p^*, where black hole formation first occurs. We will refer to the solutions with $p < p^*$ and $p > p^*$ as subcritical and supercritical, respectively. The quantity $(p - p^*)$ is a natural choice for discussing the phenomenology of the solutions, and plays a role similar to the order parameter in condensed matter systems.

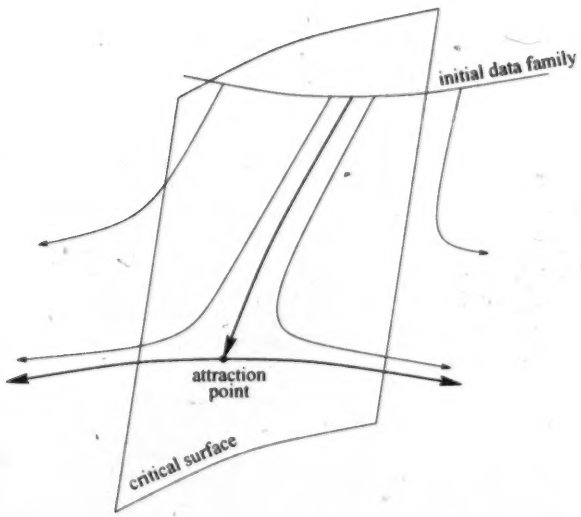


Figure 1: Schematic view of near-critical evolution of solution in the configuration space.

In retrospect, the existence of some sort of critical behavior in general relativity is not surprising, as in any system with two distinct phases there's bound to be a phase transition. However, some of its features are surprisingly simple for a system as non-linear as Einstein equations are. One could imagine that the threshold behavior would be very complicated, and could depend strongly on exactly how the parameter is tuned across threshold. It turns out, however, that this is not the case, and the critical behavior of solutions is relatively simple and, more important, *universal* in some important aspects. The characteristic properties of the near-critical gravitational collapse were first discussed by Choptuik [31]. Here we list the features which are common to critical behavior in most matter models, and postpone the discussion of the particular matter models until Section 2.5.

No mass gap

In astrophysical context, gravitational collapse normally starts from a star. The initial conditions are almost stationary, and the Chandrasekhar mass sets the scale of minimum mass for any black hole formed in the collapse. Near the threshold of black hole formation the mass $M_{\rm BH}(p)$ of the black hole produced in the collapse is discontinuous, and the height of the jump is called the mass gap. This behavior is similar to the first order phase transition in condensed matter physics.

In contrast to the above, arbitrary small black holes can be formed in the

near-critical gravitational collapse of a massless scalar field, and there is no mass gap. Critical behavior in such a system is reminiscent of the second order phase transition. Transitions with no mass gap are by far the most common in the existing literature on the critical gravitational collapse, although there are matter models that exhibit mass gaps, most notably the Yang-Mills field [32].

Mass scaling

For matter models without mass gaps, the dependence of the mass of the black hole formed in the super-critical collapse on parameter p describing the strength of gravitational interaction during the collapse turns out to be described well by a power law

$$M_{\rm BH}(p) = c_f |p - p^*|^\beta, \tag{2.10}$$

where p^* is the critical parameter value. The constant of proportionality c_f depends on the choice of the family of initial data, while the exponent β is universal for a given matter model[2].

Universality

For marginal data, both supercritical and subcritical, the evolution approaches a certain universal solution which is the same for all the families of initial data. This solution, which is unique and corresponds to the field configuration exactly at threshold of black hole formation p^*, is called the *critical solution.*

Universality of the near-critical behavior has been explained by perturbation analysis and renormalization group ideas [37, 81, 88, 67], and is rooted in the fact that critical solutions generally have only one unstable perturbation mode. In the course of evolution of the near-critical initial field configuration, all the perturbation modes contained in it decay, forgetting details of the initial data and bringing the solution closer to critical, except the single growing mode which will eventually drive the solution to black hole formation or dispersal, as illustrated in Fig. 1. In this sense, the critical solution acts as an *intermediate attractor* in the phase space of all field configurations. Because there is only a single growing mode, the codimension of the attractor is one. The eigenvalue of the growing mode determines how rapidly the solutions will eventually depart from critical, and it can be used to calculate the mass-scaling exponent β, as it is argued in Section 2.4.

[2]Initially, it was thought to be universal even across completely unrelated matter models, due to the fact that numerical calculations had shown that $\beta \approx 0.37$ for gravitational collapse of the massless scalar field [31], gravitational waves [1], and radiation fluid [37]. As no argument why these three matter models should have the same mass-scaling exponent was ever produced, it is now thought to be a coincidence, or so it would seem.

Self-similarity

The critical solution often has additional symmetry besides the spherical symmetry, called continuous or discrete self-similarity. This symmetry essentially amounts to the solution being independent of (in the case of continuous self-similarity) or periodic in (in the case of discrete self-similarity) one of the coordinates, a scale. We already encountered an example of discrete self-similarity when presenting Choptuik's results in Section 2.1. We will talk a bit more about self-similarity later in Section 2.3. The role of this symmetry in critical collapse is not completely understood.

Critical phenomena in general relativity are very interesting not only because of the simple phase transition picture above, but also because they are relevant to two fundamental problems posed before the modern theory of gravity.

The first is the cosmic censorship conjecture, the long-unsolved problem of classical general relativity (see, for example, [111, 118]). With the ability of critical solutions to produce arbitrarily small black holes and, in the critical limit, curvature singularity without an event horizon, in the course of quite generic gravitational collapse (where tuning is limited to only one parameter), they may serve as an acceptable counterexample to the cosmic censorship conjecture (see review [63] and references therein).

The second is the quantum theory of gravity. As spacetime curvature in the course of critical collapse can, in principle, reach arbitrarily large values, quantum effects will become important. The near-critical solutions could possibly be used as a test probe for the physics on Planckian scales, and relativistic quantum field theory or string theory may be studied [36]. We will say more about the possible effect of quantum corrections in Section 3.8.

Critical collapse may also be relevant to the cosmological problem of primordial black hole production [56, 83, 95, 122]. We will not discuss this topic here, and refer the reader to the papers cited above.

2.3 Self-similarity

As we mentioned above, critical solutions often are self-similar. In this section, we discuss this symmetry in more detail, give rigorous definitions of continuous and discrete self-similarity, and provide some examples.

The notion of self-similarity in physics is very old, and often emerges in various branches of physics whenever length scales of the underlying theory become unimportant for the details of the motion. In the absence of the preferred scale, it is reasonable to think that the solution will tend to look the same *on all scales*. For an extended discussion of the role of self-similarity in general relativity, see the ex-

cellent review [24]. Here we will just mention that self-similar solutions have been proposed as counterexamples to the cosmic censorship conjecture [100, 101, 110], self-similar perfect fluid solutions might be relevant in cosmology (and have now been completely classified [106, 96, 53, 54, 25, 26, 27]), and, of course, self-similar solutions play a key role in the critical gravitational collapse.

Continuous self-similarity essentially means that the solution is completely independent of one of the variables, a scale. Coordinate ambiguity in general relativity requires a more rigorous geometrical definition, however. The spacetime metric is said to be *continuously self-similar* if there exists a vector field ξ such that

$$\pounds_\xi g_{\mu\nu} = 2 g_{\mu\nu}, \tag{2.11}$$

where $\pounds_\xi$ denotes the Lie derivative. (The factor 2 in the definition is arbitrary as a vector field can always be rescaled by a constant.) The vector ξ is then a homothetic Killing vector. The matter fields should also satisfy the conditions of continuous self-similarity, which for the scalar field ϕ would be

$$\pounds_\xi \phi = 0. \tag{2.12}$$

An example of a continuously self-similar scalar field solution, which will play an important role in the discussion that follows, is due to Roberts [110]. In the null coordinates, the solution is given by

$$ds^2 = -2\,du\,dv + r^2\,d\Omega^2, \quad r^2 = u^2 - uv, \quad \phi = \frac{1}{2} \ln \left| 1 - \frac{v}{u} \right|. \tag{2.13}$$

The explicit expression for the homothetic Killing vector ξ is

$$\xi = u \frac{\partial}{\partial u} + v \frac{\partial}{\partial v}. \tag{2.14}$$

The discrete self-similarity is a weaker version of the symmetry. It basically means that the solution is a periodic, rather than a constant, function of scale. We already encountered discrete self-similarity when we discussed Choptuik's results in Section 2.1. Recall that under rescaling of space and time variables

$$t \mapsto e^{-\Delta} t, \quad r \mapsto e^{-\Delta} r \tag{2.15}$$

by a particular factor Δ, the metric and scalar field scaled as

$$ds^2 \mapsto e^{-2\Delta} ds^2, \quad \phi \mapsto \phi. \tag{2.16}$$

The rigorous definition of discrete self-similarity can be given in a coordinate-invariant way similar to (2.11) above, but for discrete displacement map, rather than infinitesimal displacement flow. The spacetime metric is said to be *discretely*

self-similar if there exists a diffeomorphism ψ on the spacetime manifold and a constant $\Delta \in \mathbb{R}$ such that

$$(\psi^*)^n g_{\mu\nu} = \exp(2\Delta n) g_{\mu\nu}, \quad \forall n \in \mathbb{Z}, \tag{2.17}$$

where ψ^* is the pullback of ψ. The matter fields should also satisfy corresponding conditions. For the scalar field, it is

$$(\psi^*)^n \phi = \phi, \quad \forall n \in \mathbb{Z}. \tag{2.18}$$

2.4 Linear perturbation analysis and mass scaling

Linear perturbation analysis is a powerful tool that allows one to calculate the mass-scaling exponent β as well as analyze the stability of a critical solution. Due to the importance of this method for our study of critical behavior in the scalar field collapse, we reproduce the general argument of Koike, Hara, and Adachi [81] in this section.

A spherically symmetric continuously self-similar solution can be written in terms of scaling coordinates which are adopted to the self-similarity of the solution, with coordinate x giving the spatial dependence of the solution, and coordinate s setting the overall scale via an e^{-s} factor[3]. Let $h = (h_1, \ldots, h_m)$ be functions of s and x which describe the solution, and satisfy a system of partial differential equations (which can be taken to be first-order without loss of generality, as that always can be arranged by introducing derivatives of the original variables as auxiliary variables)

$$\mathcal{L}\left(h, \frac{\partial h}{\partial s}, \frac{\partial h}{\partial x}\right) = 0, \tag{2.19}$$

In terms of scaling coordinates, the self-similarity of the solution is manifest in invariance of the system of partial differential equations (2.19) under the scaling transformation

$$h(s, x) \mapsto h(s + s_0, x), \tag{2.20}$$

where $s_0 \in \mathbb{R}$. A renormalization group transformation $\hat{\mathcal{R}}_{s_0}$ is a transformation on a space of functions of x

$$\hat{\mathcal{R}}_{s_0} : H \mapsto H^{(s_0)} \tag{2.21}$$

[3]For instance, for Choptuik's solution (2.1) one can take $s = -\ln(-t)$, $x = \ln(-r/t)$.

which evolves the initial data H at $s = 0$ by the partial differential equations (2.19) to $s = s_0$

$$H(x) = h(0, x), \qquad H^{(s_0)}(x) = h(s_0, x). \tag{2.22}$$

$\hat{\mathcal{R}}_{s_0}$ forms a semigroup with parameter s_0 and generator $\widehat{D\mathcal{R}} = \lim_{s_0 \to 0}(\hat{\mathcal{R}}_{s_0} - \mathbf{1})/s_0$. In this context, a self-similar solution $h_{\rm ss}(s, x) = H_{\rm ss}(x)$ can be considered as a fixed point of $\hat{\mathcal{R}}_{s_0}$ for any s_0, and satisfies $\hat{\mathcal{R}}_{s_0} H_{\rm ss} = H_{\rm ss}$, or $\widehat{D\mathcal{R}} H_{\rm ss} = 0$.

The tangent map of $\hat{\mathcal{R}}_s$ at a fixed point $H_{\rm ss}$ is a transformation on functions of x defined by

$$\hat{T}_{s_0} F = \lim_{\epsilon \to 0} \frac{\hat{\mathcal{R}}_{s_0}(H_{\rm ss} + \epsilon F) - H_{\rm ss}}{\epsilon}. \tag{2.23}$$

An eigenmode $F(x)$ of $\widehat{DT} = \lim_{s_0 \to 0}(\hat{T}_{s_0} - \mathbf{1})/s_0$ is a function that satisfies

$$\widehat{DT} F = \kappa F, \tag{2.24}$$

where $\kappa \in \mathbb{C}$. These modes determine the flow of the renormalization group near the fixed point, and are classified by $\mathrm{Re}\,\kappa$. Relevant modes ($\mathrm{Re}\,\kappa > 0$) diverge from $H_{\rm ss}$, irrelevant modes ($\mathrm{Re}\,\kappa < 0$) converge to it, and $\mathrm{Re}\,\kappa = 0$ modes are called marginal.

Let $H_{\rm c}$ be the initial data for critical parameter p^*, which eventually evolves to the fixed point

$$\lim_{s \to \infty} |H_{\rm c}^{(s)}(x) - H_{\rm ss}(x)| = 0. \tag{2.25}$$

Now consider the fate of initial datum $H_{\rm init}$ in the one-parameter family described by the parameter p, which is close to $H_{\rm c}$

$$H_{\rm init}(x) = H_{\rm c}(x) + \epsilon F(x), \tag{2.26}$$

where $\epsilon = p - p^*$. It first converges to $H_{\rm ss}$, but then eventually diverges away from the critical solution (to form a black hole or dissipate). This is schematically shown in Figure 1. If one evolves the data to some $s = s_0$, to be chosen later, using linear perturbations one can write

$$H_{\rm init}^{(s_0)} = \hat{\mathcal{R}}_{s_0} H_{\rm init} = \hat{\mathcal{R}}_{s_0}(H_{\rm c} + \epsilon F) = H_{\rm c}^{(s_0)} + \epsilon \hat{T}_{s_0} F + O(\epsilon^2). \tag{2.27}$$

In the second term, only the relevant mode with largest κ survives

$$\hat{T}_{s_0} F = \exp(s_0 \widehat{DT}) F \simeq e^{\kappa s_0} F_{\rm rel}, \tag{2.28}$$

where $F_{\rm rel}$ is the component of the relevant mode in F. Since $H_{\rm c}$ eventually converges to $H_{\rm ss}$, as shown by equation (2.25), we will have

$$H_{\rm init}^{(s_0)}(x) \simeq H_{\rm ss}(x) + \epsilon e^{\kappa s_0} F_{\rm rel}(x) \tag{2.29}$$

for large enough s_0 and $x \lesssim s_0$. When the first and the second term in (2.29) become comparable, i.e. when

$$|\epsilon| e^{(\mathrm{Re}\,\kappa) s_0} = O(1), \tag{2.30}$$

the data $H_{\mathrm{init}}^{(s_0)}$ differs from H_{ss} so much that one can tell whether the black hole will be formed, and determine its parameters. The radius of the apparent horizon (and so the mass of the black hole) will be $O(1)$ measured in x, meaning $r_{\mathrm{AH}} = O(e^{-s_0})$. Therefore, using equation (2.30), the mass of the black hole produced in the near-critical collapse is

$$M_{\mathrm{BH}} = O(e^{-s_0}) = O(|\epsilon|^{1/(\mathrm{Re}\,\kappa)}) \tag{2.31}$$

in the linear approximation. This is precisely the mass-scaling relationship (2.10), with exponent being

$$\beta = \frac{1}{\mathrm{Re}\,\kappa}. \tag{2.32}$$

Hara, Koike, and Adachi further argue [67] that this value of the mass-scaling exponent is actually exact, not just a linear approximation.

Thus the analysis of the stability of a critical solution and the calculation of the mass-scaling exponent β reduces to the study of eigenvalues of the partial differential evolution equations, given by (2.24).

2.5 Literature overview

Since Choptuik's original publication [31] appeared in 1993, a lot of work has been done in the area of the critical gravitational collapse, and a great number of publications have appeared. In this section, we give a brief overview of the available literature. For more detailed discussion, see one of the review articles [11, 21, 34, 63].

The massless scalar field is one of the simplest matter fields in general relativity, so it is not surprising that a large portion of published studies of critical behavior are based on a scalar field matter model and its various generalizations. Indeed, critical phenomena were first noticed by Choptuik [31] precisely in the model of minimally coupled massless scalar field. His results were later confirmed by independent numerical simulations [44, 65] using different coordinate systems and numerical algorithms, thus establishing that the observed effect was not a numerical artifact. The critical solution found by Choptuik has no mass gap and is discretely self-similar. It was subsequently extensively studied in Refs. [57, 58], and generalized to the charged (complex) scalar field in Refs. [59, 75].

A different family of minimally coupled massless scalar field solutions with critical behavior can be obtained by assuming continuously self-similar ansatz [110, 17, 18, 102, 103, 119, 22, 23]. Because continuous self-similarity is a much stronger symmetry than discrete self-similarity, these solutions are simpler and can be written down in a closed form (which accounts for their popularity), as opposed to Choptuik's and most other solutions known only numerically. Although these solutions are not attractors of codimension one [39], and so technically are not critical solutions in the usual sense, they provide reasonably simple toy models of critical collapse, and are studied in detail in this thesis. These solutions can be easily generalized to other scalar field couplings [40], in particular conformally coupled scalar field [98] and Brans-Dicke theory [99]. Also, the symmetry of the background allows the complete linear perturbation analysis of this model to be carried out analytically [39, 41, 42].

The generalizations of the real massless scalar field model to more exotic theories were also explored, in particular various multi-scalar sigma models [72, 73, 121]. This family includes the cases of a free complex scalar field [70, 71, 85], a real scalar field coupled to Brans-Dicke gravity [30, 33, 84] and, as a special case of the latter, an axion-dilaton model [36, 66].

The massive scalar field case is principally different from the massless one by the presence of a mass/length scale in the theory, and exhibits more complicated critical behavior. Depending on the initial conditions, both first and second order transitions can be realized [19, 28, 55]. If the initial pulse size is small compared to the Compton wavelength of the scalar field (light field), black hole formation starts from arbitrary small black holes, just as the massless case. However, if the initial pulse size is large (heavy field), black hole formation turns on at finite mass and the critical solutions are unstable soliton stars with masses of the order of the Compton wavelength of the scalar field. Such dual mode first/second order critical behavior is also encountered in the collapse of Yang-Mills field [32, 35, 60] and some other models [12, 13, 14, 86].

Although the scalar field models are the ones most extensively studied, critical phenomena were also discovered in most other matter models, in particular, in pure vacuum collapse of gravitational waves [1, 2, 15, 3], perfect fluid collapse [37, 81, 88, 68, 82, 93, 94, 116, 117], collapse of thin shells coupled to null radiation [80], collisionless matter [109] and so on. Critical behavior persists beyond spherical symmetry [61, 64, 91, 41], as well as in spacetimes of different dimensions [10, 40, 50, 79, 87, 89, 92, 104, 113, 117]. Even models with dynamics very different from Einstein gravity, such as domain walls interacting with black holes [43], seem to display some sort of critical behavior.

All this evidence suggests that the critical behavior in gravitational collapse is indeed a very general phenomenon. Large amounts of phenomenological data

have been accumulated, and some understanding has been gained, but a few mysteries still remain. Universality of the near-critical evolution and mass scaling are generally believed to be explained by the perturbation analysis and renormalization group ideas [37, 81, 88, 67], similar to those of phase transitions in condensed matter physics. However, the exact nature of the self-similarity of critical solutions and its relation to critical behavior is still largely unclear. The fact that most solutions are known only numerically presents an additional difficulty for building a complete theory of critical phenomena in general relativity, and more effort should be applied to the search of the closed form solutions or sufficiently good approximations.

The limited space and time does not allow us to go into discussion of many other interesting effects [46, 48, 62, 74], techniques [45, 47, 49], and models [78, 76, 77, 107, 108, 115]. I hope the reader will find the references compiled in the Bibliography section of this thesis useful.

Chapter 3

The Generalized Roberts Solution

In this chapter, we derive a general class of continuously self-similar, spherically symmetric scalar field solutions in spacetime of arbitrary dimension. Such solutions provide reasonably simple toy models of critical collapse, and might be relevant in the context of superstring theory, which is often said to be the next "theory of everything", as well as for understanding how critical behavior depends on the dimensionality of the spacetime. The already mentioned solution (2.13) introduced by Roberts [110, 17, 102] would be a particular case of the solutions discussed here. Some qualitative properties of the self-similar critical collapse of a scalar field in higher dimensions have been discussed in [113]. Here we aim at finding explicit closed-form solutions.

We also briefly discuss extension of minimally coupled scalar field solutions to a wider class of couplings. It is shown how different couplings of scalar field are equivalent, and several particular models are examined more closely. The procedure discussed here can be applied to any solution of Einstein-scalar field equations, although our interest lies in extending self-similar solutions.

Next we focus our attention on one particular solution of the general class presented in the beginning of this chapter, distinguished from the other minimally coupled scalar field solutions in four dimensions by the property that it lies at the threshold of the black hole formation. This solution (to which we will later refer simply as *the Roberts solution*) will serve as a background in our perturbative study of the critical collapse of the scalar field.

The material presented in this chapter mostly follows the derivation of Ref. [40].

3.1 Spherically-reduced action and field equations

Evolution of the minimally coupled scalar field in n dimensions is described by the action

$$S = \frac{1}{16\pi} \int \sqrt{-g}\, d^n x \left[R - 2(\nabla\phi)^2 \right] \tag{3.1}$$

plus surface terms. Field equations are obtained by varying this action with respect to field variables $g_{\mu\nu}$ and ϕ. However, if one is only interested in spherically symmetric solutions (as we are), it is much simpler to work with the reduced action and field equations, where this symmetry of the spacetime is factored out.

A spherically symmetric spacetime is described by the metric

$$ds^2 = d\gamma^2 + r^2 d\omega^2, \tag{3.2}$$

where $d\gamma^2 = \gamma_{AB} dx^A dx^B$ is the metric on a two-manifold, and $d\omega^2$ is the metric of a $(n-2)$-dimensional unit sphere. Essentially, the spherical symmetry reduces the number of dimensions to two, with spacetime fully described by the two-metric $d\gamma^2$ and two-scalar r. It can be shown that the reduced action describing field dynamics in spherical symmetry is

$$\begin{aligned} S_{\text{sph}} \;\propto\; & \int r^{n-2} \sqrt{-\gamma}\, d^2x \times \\ & \times \left[R[\gamma] + (n-2)(n-3) r^{-2} \left((\nabla r)^2 + 1 \right) - 2(\nabla\phi)^2 \right], \end{aligned} \tag{3.3}$$

where curvature and differential operators are calculated using the two-dimensional metric γ_{AB}; capital Latin indices run through $\{1, 2\}$, and a stroke $|$ denotes the covariant derivative with respect to the two-metric γ_{AB}. Varying the reduced action with respect to field variables γ_{AB}, r, and ϕ, we obtain Einstein-scalar field equations in spherical symmetry. After some algebraic manipulation they can be written as

$$R_{AB} - (n-2) r^{-1} r_{|AB} = 2\phi_{,A}\phi_{,B}, \tag{3.4a}$$

$$(n-3) \left[(\nabla r)^2 - 1 \right] + r \Box r = 0, \tag{3.4b}$$

$$\Box\phi + (n-2) r^{-1} \gamma^{AB} \phi_{,A} r_{,B} = 0. \tag{3.4c}$$

As usual with the scalar field, the $\Box\phi$ equation is redundant.

3.2 n-dimensional self-similar solutions

We are interested in obtaining continuously self-similar scalar field solutions in n dimensions. The geometry of the spherically symmetric spacetime is given by the two-metric γ_{AB} and the two-scalar r. A general scalar field solution can be written in double null coordinates as

$$d\gamma^2 = -2e^{-2\sigma(u,v)}du\,dv, \quad r = r(u,v), \quad \phi = \phi(u,v). \tag{3.5}$$

Imposing the continuous self-similarity conditions (2.11), the spherical reduction of which is

$$\pounds_\xi \gamma_{AB} = 2\gamma_{AB}, \quad \pounds_\xi r^2 = 2r^2, \quad \pounds_\xi \phi = 0, \tag{3.6}$$

on the metric for a homothetic Killing vector $\xi^\mu = [A(u,v), B(u,v)]$, we obtain

$$A_{,v} = 0, \quad B_{,u} = 0, \tag{3.7a}$$

$$2A\sigma_{,u} + 2B\sigma_{,v} - A_{,u} - B_{,v} = -2, \;\; Ar_{,u} + Br_{,v} = r, \;\; A\phi_{,u} + B\phi_{,v} = 0. \tag{3.7b}$$

The first two conditions (3.7a) imply that $A = A(u)$ and $B = B(v)$ are functions of one coordinate only. By a suitable definition of null coordinates, one can always put $A(u) = u$ and $B(v) = v$. Then the homothetic Killing vector is simply

$$\xi = u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v}, \tag{3.8}$$

and the remaining self-similarity conditions are

$$u\sigma_{,u} + v\sigma_{,v} = 0, \quad ur_{,u} + vr_{,v} = r, \quad u\phi_{,u} + v\phi_{,v} = 0. \tag{3.9}$$

They imply that (3.5) can be written in the form

$$d\gamma^2 = -2e^{-2\sigma(z)}du\,dv, \quad r = -u\rho(z), \quad \phi = \phi(z), \tag{3.10}$$

where dependence of metric coefficients and ϕ only on $z = -v/u$ reflects continuous self-similarity of the solution, with z being a scale-invariant variable.

We turn on the influx of the scalar field at the advanced time $v = 0$, so that the spacetime is Minkowskian to the past of this surface, and the initial conditions are specified by continuity. Signs are chosen so that $z > 0$, $\rho > 0$ in the sector of interest ($u < 0$, $v > 0$). With this metric, the Einstein-scalar equations (3.4) become

$$(n-2)\left[\rho'' z + 2\sigma'\rho - 2\sigma'\rho' z\right] = -2\rho z\phi'^2, \tag{3.11a}$$

$$2\rho(\sigma'' z + \sigma') + (n-2)\rho'' z = -2\rho z\phi'^2, \tag{3.11b}$$

$$(n-2)\left[\rho'' - 2\sigma'\rho'\right] = -2\rho\phi'^2, \tag{3.11c}$$

$$(n-3)\left[\rho'^2 z - \rho'\rho + \frac{1}{2}e^{2\sigma}\right] + \rho''\rho z = 0, \tag{3.11d}$$

$$\phi''\rho z + (n-2)\phi'\rho' z - \frac{1}{2}(n-4)\phi'\rho = 0. \tag{3.11e}$$

The prime denotes a derivative with respect to z. Combining equations (3.11a) and (3.11c), we obtain that $\sigma = \text{const}$. By appropriate rescaling of coordinates, we can always put $\sigma = 0$. Then we are left with

$$(n-2)\rho'' = -2\rho\phi'^2, \tag{3.12a}$$

$$(n-3)\left[\rho'^2 z - \rho'\rho + \frac{1}{2}\right] + \rho''\rho z = 0, \tag{3.12b}$$

$$\frac{\phi''}{\phi'} + (n-2)\frac{\rho'}{\rho} - \frac{1}{2}(n-4)z^{-1} = 0. \tag{3.12c}$$

Equation (3.12c) can be immediately integrated

$$\phi'\rho^{n-2}z^{-(n-4)/2} = c_0. \tag{3.13}$$

Substituting this result back into equation (3.12a), we get an equation for ρ only

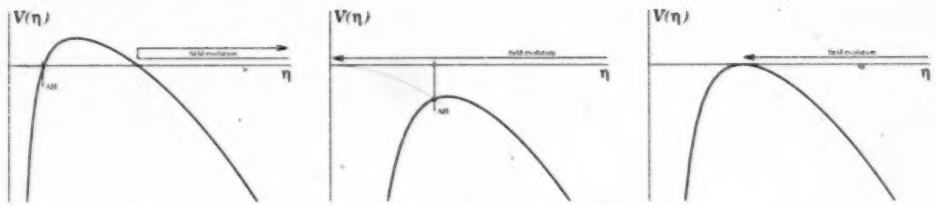
$$\rho''\rho^{2n-5} = -\frac{2c_0^2}{n-2}\,z^{n-4}. \tag{3.14}$$

It is easy to show that equation (3.14) is equivalent to equation (3.12b). This is not surprising, since the system (3.4) was redundant. For the remainder of the derivation, we will assume that $n > 3$, as the case $n = 3$ is trivial (see Section 3.4.1). Combining both equations we get a first integral of motion

$$\left[\rho'^2 z - \rho'\rho + \frac{1}{2}\right]\left(\frac{\rho^2}{z}\right)^{n-3} = \frac{2c_0^2}{(n-2)(n-3)}, \tag{3.15}$$

which contains only first derivatives of ρ, and for this reason is simpler to solve than either one of equations (3.12b) and (3.14). Equation (3.15) is a generalized homogeneous one, and can be solved by substitution

$$x = \frac{1}{2}\ln z, \quad \rho = \sqrt{z}\eta(x), \quad \rho' = \frac{1}{2}z^{-1/2}(\dot\eta + \eta), \tag{3.16}$$



(a) (b) (c)

Figure 2: Field evolution in effective potential: (a) subcritical, (b) supercritical, (c) critical.

where the dot denotes the derivative with respect to the new variable x. With this substitution, equations (3.15) and (3.13) become

$$\dot{\eta}^2 = \eta^2 - 2 + c_1 \eta^{-2(n-3)}, \tag{3.17}$$

$$\dot{\phi} = 2c_0 \eta^{-(n-2)}, \tag{3.18}$$

where we defined the constant

$$c_1 = \frac{8c_0^2}{(n-2)(n-3)} > 0. \tag{3.19}$$

The above equation (3.17) for η formally describes motion of a particle with zero energy in the potential

$$V(\eta) = 2 - \eta^2 - c_1 \eta^{-2(n-3)}, \tag{3.20}$$

so we can tell the qualitative behavior of η without actually solving equation (3.17).

Initial conditions are specified by continuous matching of the solution to the flat spacetime across the surface $v = 0$, when the scalar field influx is turned on. Since on that surface $r \neq 0$, the value of $\eta = r/\sqrt{-uv}$ starts from infinity at $x = -\infty$, and rolls towards zero. What happens next depends on the shape of the potential. If there is a region with $V(\eta) > 0$, as in Fig. 2a, η will reach a turning point and will go back to infinity as $x = \infty$. If $V(\eta) < 0$ everywhere, as in Fig. 2b, there is nothing to stop η from reaching zero, at which point a singularity is formed. Finally, if $V(\eta)$ has a second-order zero, as in Fig. 2c, η will take forever reaching it.

Of course, the variables separate, and equation (3.17) can be integrated:

$$x = \pm \int \frac{d\eta}{\sqrt{\eta^2 - 2 + c_1 \eta^{-2(n-3)}}} + c_2. \tag{3.21}$$

The plus or minus sign in front of the integral depends on the sign of the derivative of η. Initial conditions imply that initially η comes from infinity towards zero, i.e. its derivative is negative, and so we must pick the branch of the solution which started out with a minus sign. The constant c_2 corresponds to a coordinate freedom in the choice of the origin of x, while the constant c_1 is a real parameter of the solution.

Unfortunately, the integral cannot be evaluated in a closed form for arbitrary n. But if the integral is evaluated, and we can invert it to get η as a function of x, then the solution for r is obtained by using definitions (3.16) and (3.10), and the solution for ϕ is obtained by integrating relations (3.18) or (3.13). Even when the integral (3.21) cannot be evaluated, one can still write an expression for the spacetime metric in terms of elementary functions by using η and $\chi = \sqrt{-uv}$ as coordinates

$$d\gamma^2 = \frac{-2\chi^2\, d\eta^2}{\eta^2 - 2 + c_1\eta^{-2(n-3)}} + 2d\chi^2, \quad r = \chi\eta. \tag{3.22}$$

The solution for ϕ still would require integration of

$$\frac{d\phi}{d\eta} = \frac{2c_0}{\sqrt{\eta^{2(n-1)} - 2\eta^{2(n-2)} + c_1\eta^2}}. \tag{3.23}$$

3.3 Critical behavior

The one-parameter family of self-similar scalar field solutions in n dimensions we constructed above exhibits critical behavior as the parameter c_1 is tuned. In this section we investigate black hole formation in the collapse.

In spherical symmetry, existence and position of the apparent horizon are given by the vanishing of $(\nabla r)^2 = 0$, which translates to $\rho' = 0$, or $\dot{\eta} + \eta = 0$ in our notation. Therefore, at the apparent horizon we have

$$\dot{\eta}^2 - \eta^2 = c_1\eta^{-2(n-3)} - 2 = 0, \tag{3.24}$$

and the black hole is formed if the value of η reaches

$$\eta_{\rm AH}^2 = \left(\frac{c_1}{2}\right)^{\frac{1}{n-3}}. \tag{3.25}$$

As we have discussed above, depending on the value of c_1, the values of the field η either reach the turning point and return to infinity, or go all the way to zero. The critical solution separates the two cases, and is characterized by potential $V(\eta)$ having a second order zero, i.e. $V(\eta_*) = V'(\eta_*) = 0$ at some point

η_*. Differentiating expression (3.20) for the potential $V(\eta)$, we see that it has a second order zero at

$$\eta_*^2 = 2\,\frac{n-3}{n-2}, \tag{3.26}$$

if and only if the value of the constant c_1 is

$$c_1^* = \frac{1}{n-3}\left[2\,\frac{n-3}{n-2}\right]^{n-2}. \tag{3.27}$$

If the value of parameter c_1 is less than critical, $c_1 < c_1^*$, the value of η turns around at the turning point, and never reaches the point of the apparent horizon formation, which is located in the forbidden zone, as illustrated in Fig. 2a. This case is subcritical evolution of the field. If $c_1 > c_1^*$, the value of η reaches the point where the apparent horizon is formed, and proceeds to go to zero, at which place there is a singularity inside the black hole. This supercritical evolution is illustrated in Fig. 2b.

The mass of the black hole formed in the supercritical collapse is

$$M = \frac{1}{2}\,r_{\rm AH} = -\frac{1}{2}\,u\sqrt{z_{\rm AH}}\,\eta_{\rm AH}. \tag{3.28}$$

It grows infinitely if we wait long enough, and will absorb all the field influx coming from past infinity. This happens because the solution is self-similar, and creates a problem for discussing mass scaling in the near-critical collapse. Switching the influx of the scalar field off at some finite advanced time will alleviate this problem. This is discussed in more detail later, in Section 3.7.2.

3.4 Particular cases

In this section we consider several particular cases for which the general solution (3.21) is simplified. Particularly important is the $n = 4$ case.

3.4.1 $n = 3$

As we have already mentioned, for $n = 3$ the only self-similar scalar field solution of the form (3.10) is trivial. To see it, note that equation (3.12b) implies that $\rho'' = 0$ if $n = 3$, and so $\rho = \alpha z + \beta$ and $r = \alpha v - \beta u$. From equation (3.12a) it then follows that $\phi = \text{const}$. The spacetime is flat.

3.4.2 $n = 4$

Integration (3.21) can be carried out explicitly

$$\begin{aligned} x &= -\int \frac{d\eta}{\sqrt{\eta^2 - 2 + c_1\eta^{-2}}} + c_2 \\ &= -\frac{1}{2}\ln\left|\eta^2 - 1 + \sqrt{\eta^4 - 2\eta^2 + c_1}\right| + c_2, \end{aligned} \tag{3.29}$$

and the result inverted

$$\eta^2 = \frac{1}{2}e^{-2(x-c_2)} + 1 + \frac{1-c_1}{2}e^{2(x-c_2)}, \tag{3.30}$$

to give the solution in the closed form

$$\rho = \sqrt{\frac{e^{2c_2}}{2} + z + \frac{1-c_1}{2e^{2c_2}}z^2}. \tag{3.31}$$

By appropriately rescaling coordinates, we can put $e^{2c_2} = 2$. After redefining the parameter of the solution to be $p = (c_1 - 1)/4$, the solution takes on the following simple form:

$$\rho = \sqrt{1 + z - pz^2}, \quad r = \sqrt{u^2 - uv - pv^2}. \tag{3.32}$$

The scalar field ϕ is reconstructed from equation (3.13)

$$\phi' = c_0\rho^{-2} = \frac{1}{2}\frac{\sqrt{1+4p}}{1+z-pz^2}, \tag{3.33}$$

to give

$$\begin{aligned} \phi &= \operatorname{arctanh}\frac{2pz-1}{\sqrt{1+4p}} + \text{const} \\ &= \frac{1}{2}\ln\left[-\frac{2pz-1+\sqrt{1+4p}}{2pz-1-\sqrt{1+4p}}\right] + \text{const}. \end{aligned} \tag{3.34}$$

This family of solutions has been derived earlier in [110, 17, 102]. The critical parameter value is $p^* = 0$, and for $p > 0$ the black hole is formed. The solution at the threshold of black hole formation is

$$r = \sqrt{u^2 - uv}, \quad \phi = \frac{1}{2}\ln\left[1 - \frac{v}{u}\right]. \tag{3.35}$$

3.4.3 $n = 5, 6$

The integral (3.21) can be written in terms of elliptic functions, which becomes apparent with the change of variable $\bar{\eta} = \eta^{-2}$

$$x = \frac{1}{2} \int \frac{d\bar{\eta}}{\bar{\eta}\sqrt{1 - 2\bar{\eta} + c_1 \bar{\eta}^{n-2}}}, \tag{3.36}$$

and the solution $\eta(x)$ is given implicitly. However, integrals corresponding to critical solutions simplify, and can be given in terms of elementary functions for $n = 5, 6$. It happens because the potential factors since it has a second order zero, therefore reducing the power of η in the radical by two. The results of integration for critical solutions are

$$\begin{aligned} n = 5: \quad x &= -\int \frac{\eta^2 d\eta}{(\eta^2 - \frac{4}{3})\sqrt{\eta^2 + \frac{2}{3}}} \\ &= -\operatorname{arcsinh}\left(\sqrt{\frac{3}{2}}\eta\right) \\ &\quad + \frac{1}{\sqrt{6}} \operatorname{arctanh}\left(\frac{\sqrt{3}\eta + 1}{\sqrt{\frac{9}{2}\eta^2 + 3}}\right) \\ &\quad + \frac{1}{\sqrt{6}} \operatorname{arctanh}\left(\frac{\sqrt{3}\eta - 1}{\sqrt{\frac{9}{2}\eta^2 + 3}}\right); \end{aligned} \tag{3.37}$$

$$\begin{aligned} n = 6: \quad x &= -\int \frac{\eta^3 d\eta}{(\eta^2 - \frac{3}{2})\sqrt{\eta^4 + \eta^2 + \frac{3}{4}}} \\ &= -\frac{1}{2} \operatorname{arcsinh}\left(\sqrt{2}(\eta^2 + \frac{1}{2})\right) \\ &\quad + \frac{1}{2\sqrt{2}} \operatorname{arctanh}\left(\frac{1}{\sqrt{2}} \frac{\frac{4}{3}\eta^2 + 1}{\sqrt{\eta^4 + \eta^2 + \frac{3}{4}}}\right). \end{aligned} \tag{3.38}$$

The dependence $\eta(x)$ is still given implicitly. The critical value of the parameter c_1^* is $32/27$ for $n = 5$, and $27/16$ for $n = 6$.

3.4.4 Higher dimensions

For higher dimensions, the integral (3.21) cannot be taken in terms of elementary functions. One can write the metric in the closed form (3.22), but the expression

for scalar field (3.23) would remain a quadrature. One can easily obtain the answer by numerical integration, however.

3.5 General scalar field coupling

In this section we discuss in detail how solutions of the minimally coupled scalar field model can be generalized to a much wider class of couplings. The fact that essentially all couplings of a free scalar field to its kinetic term and scalar curvature are equivalent has been long used [9] to study scalar field models with non-minimal coupling. In particular, this idea has been applied to extend the four-dimensional Roberts solution to conformal coupling [98] and Brans-Dicke theory [99].

3.5.1 Equivalence of couplings

Suppose that the action describing evolution of the scalar field in n-dimensional spacetime is

$$S = \frac{1}{16\pi} \int \sqrt{-g}\, d^n x \left[F(\phi) R - G(\phi) (\nabla\phi)^2 \right] \tag{3.39}$$

plus surface terms, where the couplings F and G are smooth functions of field ϕ. Also suppose that the signs of couplings F and G correspond to the case of gravitational attraction. We will demonstrate that this action reduces to the minimally coupled one by redefinition of the fields $g_{\mu\nu}$ and ϕ. First, let us introduce a new metric $\hat{g}_{\mu\nu}$ that is related to the old one by the conformal transformation

$$\hat{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}, \quad \hat{g}^{\mu\nu} = \Omega^{-2} g^{\mu\nu}, \quad \sqrt{-\hat{g}} = \Omega^n \sqrt{-g}, \tag{3.40}$$

and denote quantities and operators calculated using $\hat{g}_{\mu\nu}$ by a hat. Scalar curvatures calculated using old and new metrics are related

$$R = \Omega^2 \hat{R} + 2(n-1)\Omega\hat{\Box}\Omega - n(n-1)(\hat{\nabla}\Omega)^2, \tag{3.41}$$

as are field gradients

$$(\nabla\phi)^2 = \Omega^2 (\hat{\nabla}\phi)^2, \tag{3.42}$$

Writing the action (3.39) in terms of the metric $\hat{g}_{\mu\nu}$, we obtain

$$\begin{aligned} S &= \frac{1}{16\pi} \int \Omega^{-n} \sqrt{-\hat{g}}\, d^n x \times \\ &\times \left[F\{\Omega^2 \hat{R} + 2(n-1)\Omega\hat{\Box}\Omega - n(n-1)(\hat{\nabla}\Omega)^2\} - G\Omega^2 (\hat{\nabla}\phi)^2 \right]. \end{aligned} \tag{3.43}$$

By choosing the conformal factor to be

$$\Omega^{n-2} = F, \tag{3.44}$$

the factor in front of the curvature $\hat{R}$ can be set to one. Substitution of the definition of Ω into the above action, and integration by parts of the $\hat{\Box}$ operator yields

$$S = \frac{1}{16\pi} \int \sqrt{-\hat{g}}\, d^n x \left[\hat{R} - \left(\frac{G}{F} + \frac{n-1}{n-2} \frac{F'^2}{F^2} \right) (\hat{\nabla}\phi)^2 \right]. \tag{3.45}$$

The kinetic term in action (3.45) can be brought into minimal form by introduction of a new scalar field $\hat{\phi}$, related to the old one by

$$2(\hat{\nabla}\hat{\phi})^2 = \left(\frac{G}{F} + \frac{n-1}{n-2} \frac{F'^2}{F^2} \right) (\hat{\nabla}\phi)^2. \tag{3.46}$$

Thus, we have shown that with field redefinitions

$$\hat{\phi} = \frac{1}{\sqrt{2}} \int \left(\frac{G}{F} + \frac{n-1}{n-2} \frac{F'^2}{F^2} \right)^{\frac{1}{2}} d\phi, \quad \hat{g}_{\mu\nu} = F^{2/(n-2)} g_{\mu\nu}, \tag{3.47}$$

the generally coupled scalar field action (3.39) becomes minimally coupled

$$S = \frac{1}{16\pi} \int \sqrt{-\hat{g}}\, d^n x \left[\hat{R} - 2(\hat{\nabla}\hat{\phi})^2 \right]. \tag{3.48}$$

This equivalence allows one to construct solutions of the model with general coupling (3.39) from the solutions of the minimally coupled scalar field by means of the inverse of relation (3.47), provided the said inverse is well-defined. However, there may be some restrictions on the range of ϕ so that field redefinitions give real $\hat{\phi}$ and positive-definite metric $\hat{g}_{\mu\nu}$. This means that not all the branches of the solution in general coupling may be covered by translating the minimally coupled solution. Technically speaking, the correspondence between solutions of the minimally coupled theory and the generally coupled theory is one-to-one where defined, but not onto.

However, one has to be careful making claims about global structure and critical behavior of the generalized solutions based solely on the properties of the minimally-coupled solution. The scalar field solutions encountered in critical collapse often lead to singular conformal transformations, which could, in principle, change the structure of spacetime. In the simplest case of non-singular conformal transformation (i.e. when coupling F is bounded and a lower bound is greater than zero) global properties of the minimally coupled solution are preserved, and all important features of near-critical collapse carry over on the generalized solution unchanged.

3.5.2 Examples

To illustrate the discussion above, we consider two often used scalar field couplings as examples. They are non-minimal coupling and dilaton gravity.

Conformal coupling

Non-minimally coupled scalar field in n dimensions is described by action

$$S = \frac{1}{16\pi}\int \sqrt{-g}\, d^n x \left[(1-2\xi\phi^2)R - 2(\nabla\phi)^2\right], \tag{3.49}$$

where ξ is the coupling parameter. Coupling factors are $F = 1-2\xi\phi^2$, $G = 2$ and so the field redefinition (3.47) looks like

$$\begin{aligned}\hat{\phi} &= \int \frac{\sqrt{1-2\xi\phi^2+2\xi_c^{-1}\xi^2\phi^2}}{1-2\xi\phi^2}\, d\phi \\ &= \frac{1}{\sqrt{2}}\sqrt{\xi^{-1}-\xi_c^{-1}}\,\arcsin\left[\sqrt{2}\sqrt{\xi-\xi_c^{-1}\xi^2}\,\phi\right] \\ &\quad + \frac{1}{\sqrt{2\xi_c}}\,\mathrm{arcsinh}\left[\frac{\sqrt{2\xi_c^{-1}}\,\xi\phi}{\sqrt{1-2\xi\phi^2}}\right],\end{aligned} \tag{3.50}$$

where

$$\xi_c = \frac{1}{4}\frac{n-2}{n-1}. \tag{3.51}$$

Particularly interesting is the case of conformal coupling $\xi = \xi_c$, because the field redefinition

$$\hat{\phi} = \frac{1}{\sqrt{2\xi_c}}\,\mathrm{arctanh}\left[\sqrt{2\xi_c}\,\phi\right] \tag{3.52}$$

can be inverted explicitly to give the recipe for obtaining conformally coupled solutions from minimally coupled ones. It is

$$\phi = \frac{1}{\sqrt{2\xi_c}}\tanh\left[\sqrt{2\xi_c}\,\hat{\phi}\right], \tag{3.53}$$

$$g_{\mu\nu} = \frac{\hat{g}_{\mu\nu}}{(1-2\xi_c\phi^2)^{2/(n-2)}} = \cosh^{4/(n-2)}\left[\sqrt{2\xi_c}\,\hat{\phi}\right]\hat{g}_{\mu\nu}. \tag{3.54}$$

In particular, the four dimensional self-similar scalar field solution becomes

$$\phi = \sqrt{3}\tanh\left[\frac{1}{\sqrt{3}}\,\mathrm{arctanh}\,\frac{2pz-1}{\sqrt{1+4p}}\right], \tag{3.55}$$

$$ds^2 = \cosh^2\left[\frac{1}{\sqrt{3}}\,\mathrm{arctanh}\,\frac{2pz-1}{\sqrt{1+4p}}\right]\times \times\left\{-2du\,dv+(u^2-uv-pv^2)\,d\omega^2\right\}. \tag{3.56}$$

in the conformally coupled model. This last expression was considered in [98].

Dilaton gravity

Another useful example is dilaton gravity described by the action

$$S=\frac{1}{16\pi}\int\sqrt{-g}\,d^nx\,e^{-2\phi}\left[R+4(\nabla\phi)^2\right]. \tag{3.57}$$

Substituting coupling factors $F=e^{-2\phi}$, $G=-4e^{-2\phi}$ into the relationship (3.47), one can see that the scalar field redefinition is a simple scaling

$$\hat{\phi}=\sqrt{\frac{2}{n-2}}\,\phi,\quad \phi=\sqrt{\frac{n-2}{2}}\,\hat{\phi}, \tag{3.58}$$

and the metrics differ by an exponential factor only

$$g_{\mu\nu}=\exp\left[\sqrt{\frac{2}{n-2}}\,2\hat{\phi}\right]\hat{g}_{\mu\nu}. \tag{3.59}$$

In particular, the four dimensional self-similar scalar field solution becomes

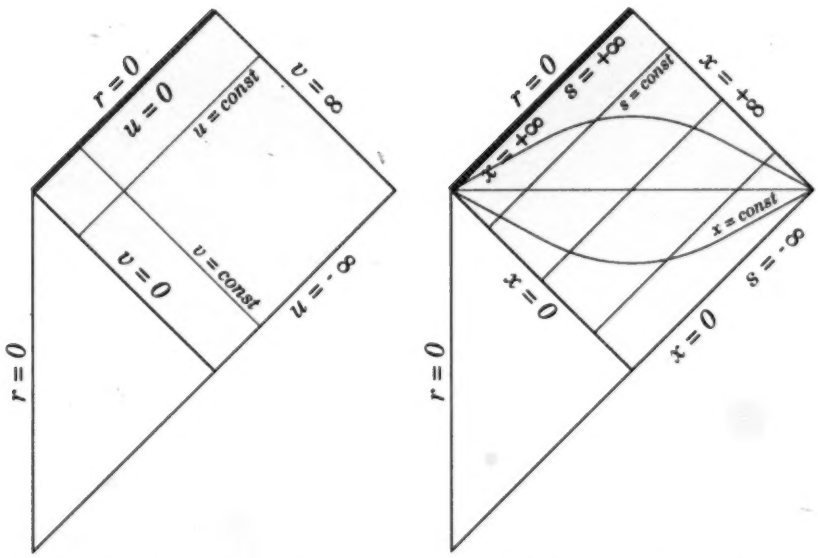
$$\phi=\mathrm{arctanh}\,\frac{2pz-1}{\sqrt{1+4p}}, \tag{3.60}$$

$$ds^2=e^{2\phi}\left\{-2du\,dv+(u^2-uv-pv^2)\,d\omega^2\right\}. \tag{3.61}$$

in dilaton gravity.

3.6 Properties of the Roberts solution

In this section we consider one particular solution from the general class of self-similar solutions, distinguished from the other minimally coupled scalar field solutions in four dimensions by the property that it lies at the threshold of the black hole formation. Since it plays a key role by being a background in our perturbative study of the critical collapse of the scalar field, more detailed examination of its properties is in order.



(null coordinates) (scaling coordinates)

Figure 3: Global structure of the Roberts solution: The scalar field influx is turned on at $v = 0$; spacetime is flat before that. The field evolution occurs in the shaded region of the diagram, and there is a null singularity in the center of the spacetime.

3.6.1 Global structure

As we have shown above in Section 3.4, the four dimensional continuously self-similar spherically symmetric scalar field solution at the threshold of the black hole formation is given by the metric

$$ds^2 = -2\,du\,dv + r^2\,d\Omega^2, \tag{3.62}$$

with

$$r = \sqrt{u^2 - uv}, \quad \phi = \frac{1}{2}\ln\left[1 - \frac{v}{u}\right]. \tag{3.63}$$

The global structure of the corresponding spacetime is shown in Fig. 3. The influx of the scalar field is turned on at the advanced time $v = 0$, so that the Roberts spacetime is smoothly matched to Minkowskian flat spacetime to the past of this surface. The junction conditions, required for continuity of the solution there, serve as boundary conditions for the field equations. We discuss them in detail in Section 3.7.

The spacetime has a null singularity in the center, with the curvature invariants diverging at $u = 0$ as

$$R = \frac{v}{u(v-u)^2}, \quad R_{\mu\nu}R^{\mu\nu} = R^2, \quad C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta} = \frac{4}{3}R^2. \tag{3.64}$$

The scalar field flux is given by its stress-energy tensor, which has the following non-trivial components

$$8\pi T_{uu} = \frac{1}{2}\frac{v^2}{u^2(v-u)^2}, \quad 8\pi T_{vv} = \frac{1}{2}\frac{1}{(v-u)^2}, \quad 8\pi T_{\theta\theta} = \frac{1}{2}\frac{v}{v-u}. \tag{3.65}$$

The remaining gravitational degrees of freedom are described by the Weyl tensor, which has the following non-vanishing components

$$-C^{uv}{}_{uv} = -C^{\theta\phi}{}_{\theta\phi} = 2C^{u\theta}{}_{u\theta} = 2C^{u\phi}{}_{u\phi} = 2C^{v\theta}{}_{v\theta} = 2C^{v\phi}{}_{v\phi} = \frac{1}{3}R. \tag{3.66}$$

3.6.2 Newman-Penrose formalism

It is instructive to also write down the Roberts solution (3.62,3.63) in Newman-Penrose formalism. In spherically symmetric spacetimes, it is natural to take the tetrad vectors based on the radial null ray directions. In null coordinates, its components are given by

$$l_\mu = (1,0,0,0), \tag{3.67a}$$

$$n_\mu = (0,1,0,0), \tag{3.67b}$$

$$m_\mu = \frac{1}{\sqrt{2}}(0,0,r,ir\sin\theta). \tag{3.67c}$$

With this tetrad choice, various Newman-Penrose quantities can be easily calculated for the Roberts solution. (We must note though, that we provide explicit expressions for them in original Newman-Penrose sign convention, which differs from the usual Misner-Thorne-Wheeler sign convention used throughout this thesis.) One finds that the spin coefficients are nearly all zero, except for

$$\rho = \frac{1}{2}\frac{u}{r^2}, \quad \mu = \frac{1}{2}\frac{2u-v}{r^2}, \quad \alpha = \frac{\sqrt{2}}{4}\frac{\cot\theta}{r} = -\beta. \tag{3.68}$$

The non-vanishing Ricci scalars are

$$\Phi_{00} = \frac{u}{v}\frac{R}{4}, \quad \Phi_{11} = -\frac{R}{8}, \quad \Phi_{22} = \frac{v}{u}\frac{R}{4}, \quad \Lambda = \frac{R}{24}. \tag{3.69}$$

Finally, the only non-vanishing Weyl scalar is

$$\Psi_2 = \frac{R}{6}. \tag{3.70}$$

The Roberts solution is algebraically special, and is classified as Petrov type D. The high degree of symmetry of the background indicated by the vanishing of most Newman-Penrose quantities might be responsible for the decoupling of perturbation equations which we will encounter in the next chapter.

3.6.3 Scaling coordinates

For the calculations in the following chapters, it is advantageous to introduce a new coordinate system adapted to the self-similarity, because then the scale-invariance of the solution becomes apparent. Such a coordinate system, which we will call scaling coordinates, can be implemented by

$$x = \frac{1}{2}\ln\left[1 - \frac{v}{u}\right], \quad s = -\ln(-u), \tag{3.71}$$

with the inverse transformation

$$u = -e^{-s}, \quad v = e^{-s}(e^{2x} - 1). \tag{3.72}$$

The signs are chosen to make the arguments of the logarithm positive in the region of interest ($v > 0$, $u < 0$), where the field evolution occurs. In these coordinates the metric (3.62) becomes

$$g_{\mu\nu}\,dx^\mu dx^\nu = 2e^{2(x-s)}\left[(1 - e^{-2x})ds^2 - 2ds\,dx\right] + r^2\,d\Omega^2, \tag{3.73}$$

and the Roberts solution (3.63) is simply

$$r = e^{x-s}, \quad \phi = x. \tag{3.74}$$

Observe that the scalar field ϕ does not depend on the scale variable s at all, and the only dependence of the metric coefficients on the scale is through the conformal factor e^{-2s}. This is a direct expression of the geometric definition of continuous self-similarity as the existence of a homothetic Killing vector ξ such that

$$\pounds_\xi g_{\mu\nu} = 2g_{\mu\nu}, \quad \pounds_\xi \phi = 0, \tag{3.75}$$

where $\pounds$ denotes the Lie derivative. In scaling coordinates, the homothetic Killing vector is simply

$$\xi = -\frac{\partial}{\partial s}. \tag{3.76}$$

As an interesting side observation, we note that the two-dimensional part of the metric (3.73) is related to the two-dimensional black hole metric

$$d\tilde{\gamma}^2 = (1 - e^{-2x})ds^2 - 2ds\,dx \tag{3.77}$$

by a conformal transformation $d\gamma^2 = 2e^{2(x-s)}d\tilde{\gamma}^2$. This transformation maps the homothetic Killing vector ξ into a Killing vector of the two-dimensional black hole spacetime. It is spacelike in the region ($v > 0$, $u < 0$) where the field evolution occurs, hence the Roberts solution is mapped into the *interior* of the two-dimensional black hole by this transformation.

3.6.4 Curvature coordinates

In Choptuik's work on scalar field collapse [31], as well as in some other numerical studies, diagonal Schwarzschild-like coordinates are used to describe the spacetime. In order to relate our results to these works, we will need to derive the expression for the Roberts metric in diagonal form

$$ds^2 = -\alpha\, dt^2 + \beta\, dr^2 + r^2\, d\Omega^2. \tag{3.78}$$

One can straightforwardly check that the coordinate change

$$t = -\exp\left[-s - \frac{1}{2}\, e^{2x}\right], \quad r = \exp[x - s] \tag{3.79}$$

diagonalizes the Roberts metric (3.73). By self-similarity, the quantity t/r, as well as the metric coefficients α and β do not depend on the scale s, but only on the coordinate x. The metric coefficients, written as functions of x, are

$$\alpha = 2\,\frac{\exp[e^{2x}]}{1 + e^{2x}}, \quad \beta = 2\,\frac{1}{1 + e^{-2x}}. \tag{3.80}$$

If one wishes, one can rewrite them as explicit functions of t/r, using

$$x = \frac{1}{2}\,\ln W(r^2/t^2), \tag{3.81}$$

in terms of Lambert's W-function, which is defined by the solution of transcendental equation

$$W \exp(W) = x. \tag{3.82}$$

The expressions for metric coefficients are then

$$\alpha = 2\,\frac{\exp[W(r^2/t^2)]}{1 + W(r^2/t^2)}, \quad \beta = 2\,\frac{W(r^2/t^2)}{1 + W(r^2/t^2)}. \tag{3.83}$$

However, the coefficients α and β cannot be written in closed form in terms of elementary functions of t/r.

As you can see from expressions for the metric above, diagonal Schwarzschild coordinates are not particularly well-suited for the description of the Roberts spacetime. On top of the complicated metric form, one artifact of the diagonal coordinate system is that the null singularity at $u = 0$ gets compressed into a point $r = t = 0$. Also, slices $t = \text{const}$ cut across the $v = 0$ hypersurface, so one has to be careful with possible discontinuities of the solution there.

3.7 Switching the scalar field influx on and off

We conclude our discussion of the Roberts solution by considering in detail how one can turn the influx of the scalar field on and off. Technically, this is achieved by gluing the Robert solution to appropriate spacetimes across the null hypersurface of constant advanced time [119]: to flat Minkowski spacetime before $v = 0$ when switching the field on, and to an outgoing Vaidya solution after $v = v_0$ when switching the field off. In the next two sections, we derive the junction conditions for these two cases, following the general formalism of thin null shells by Barrabès and Israel [7].

3.7.1 Matching with flat spacetime

Consider the general spherically symmetric metric in null coordinates

$$ds^2 = -2e^{2\sigma}\,du\,dv + r^2\,d\Omega^2. \tag{3.84}$$

To fix the geometry of the soldering of spherically symmetric spacetimes uniquely, one must match radial functions across the constant advanced time hypersurface. To this end, we rewrite the above metric in terms of advanced Eddington coordinates on both sides

$$ds^2 = -e^{\psi}\,dv\,(fe^{\psi}\,dv - 2\,dr) + r^2\,d\Omega^2. \tag{3.85}$$

The metric coefficients in Eddington coordinates can be calculated from those in null coordinates by

$$e^{\psi} = -\frac{e^{2\sigma}}{r_{,u}}, \quad f \equiv (\nabla r)^2 = -2\,\frac{r_{,u}r_{,v}}{e^{2\sigma}}. \tag{3.86}$$

The surface density and pressure of the null shell are then determined by jumps of the metric coefficients across the $v = \text{const}$ surface

$$4\pi r^2\epsilon = [m], \quad 8\pi P = [\psi_{,r}], \tag{3.87}$$

where the local mass function $m(v, r)$ is introduced, as usual, by

$$f = 1 - \frac{2m}{r}. \tag{3.88}$$

For Minkowski spacetime $f = 1$ and $\psi = \text{const}$, so below the $v = 0$ hypersurface surface we have

$$m(v < 0) = 0, \quad \psi_{,r}(v < 0) = 0. \tag{3.89}$$

For the Roberts solution (3.62,3.63), we have $\sigma = 0$, $r^2 = u^2 - uv$, so a direct calculation gives

$$m = -\frac{uv}{4r}, \quad \psi_{,r} = \frac{1}{r}\frac{v^2}{4r^2 + v^2}. \tag{3.90}$$

Obviously,

$$\lim_{v \to +0} m = 0, \quad \lim_{v \to +0} \psi_{,r} = 0, \tag{3.91}$$

so the Roberts solution is indeed attached smoothly to the flat spacetime, without a delta function-like stress-energy tensor singularity associated with a massive null shell.

3.7.2 Matching with outgoing Vaidya solution

If the incoming flux of scalar field is switched off at some finite advanced time v_0, by causality, this won't affect the solution at earlier times. However, even when the incoming flux had stopped, there still will be some outgoing flux leaking out to $\mathcal{J}^+$. As we will show, the solution in this regime is described by an outgoing Vaidya solution, given in coordinates (U, r, θ, ϕ) by the metric

$$ds^2 = -f\,dU^2 - 2\,dU\,dr + r^2\,d\Omega^2, \quad f = 1 - \frac{2m(U)}{r}. \tag{3.92}$$

Since outgoing Vaidya solution cannot be readily written in terms of advanced Eddington coordinates, we have to employ a more general formalism to derive junction conditions than in the last section. To match the outgoing Vaidya solution across the null surface Σ defined by $v = \text{const}$, or in differential form, by

$$f\,dU + 2dr = 0, \tag{3.93}$$

following [7], we calculate the holonomic basis vectors tangent to Σ to be

$$e^{\mu}_{(r)} = \left(-\frac{2}{f}, 1, 0, 0\right), \tag{3.94a}$$

$$e^{\mu}_{(\theta)} = (0, 0, 1, 0), \tag{3.94b}$$

$$e^{\mu}_{(\phi)} = (0, 0, 0, 1). \tag{3.94c}$$

The normal n and transversal N vectors to the surface Σ are

$$n^{\mu} = -e^{\mu}_{(r)}, \quad N^{\mu} = \left(0, \frac{f}{2}, 0, 0\right). \tag{3.95}$$

They satisfy the following conditions

$$N \cdot N = 0, \quad N \cdot n = -1, \quad N_a = (1, 0, 0). \tag{3.96}$$

The transverse extrinsic curvature, defined by

$$\mathcal{K}_{ab} = -N_\mu e^\nu_{(b)} \nabla_\nu e^\mu_{(a)}, \tag{3.97}$$

has non-vanishing components

$$\mathcal{K}_{rr} = \frac{4}{r} \frac{m'(U)}{f^2}, \quad \mathcal{K}_{\theta\theta} = \frac{r}{2} f. \tag{3.98}$$

From the Roberts solution side, $r^2 = u^2 - uv$, so the r holonomic basis vector tangent to Σ is

$$e^\mu_{(r)} = \left(\frac{2r}{2u - v}, 0, 0, 0 \right), \tag{3.99}$$

with the other two vectors being the same as (3.94) because both solutions are spherically symmetric. The normal n and transversal N vectors to the surface Σ are

$$n^\mu = -e^\mu_{(r)}, \quad N^\mu = \left(0, -\frac{2u - v}{2r}, 0, 0 \right), \tag{3.100}$$

and they also satisfy the conditions (3.96). The transverse extrinsic curvature from the Roberts side has non-vanishing components

$$\mathcal{K}_{rr} = -\frac{v^2}{r(2u - v)^2} = -\frac{1}{r} \frac{v^2}{4r^2 + v^2}, \tag{3.101a}$$

$$\mathcal{K}_{\theta\theta} = \frac{1}{4} \frac{(2u - v)u}{r} = \frac{2r^2 + uv}{4r}. \tag{3.101b}$$

The two solutions would match smoothly across the surface Σ if and only if the transverse vectors and transverse extrinsic curvatures are continuous across that surface, i.e. their jumps are zero

$$[N \cdot N] = 0, \quad [Na] = 0, \quad [\mathcal{K}_{ab}] = 0. \tag{3.102}$$

In our case, the first two conditions are satisfied automatically, while the condition on $\mathcal{K}_{rr}$ gives

$$[\mathcal{K}_{rr}] = \frac{4}{r} \frac{m'(U)}{f^2} + \frac{1}{r} \frac{v_0^2}{4r^2 + v_0^2} = 0. \tag{3.103}$$

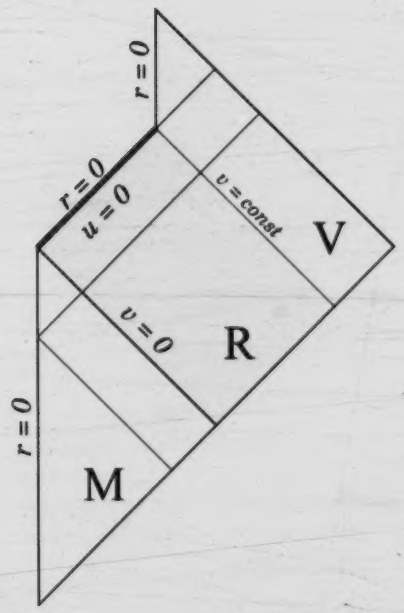


Figure 4: Global structure of the Roberts solution matched to the outgoing Vaidya solution: The scalar field influx is turned on at $v = 0$, and switched off again at $v = v_0$.

If we denote the on-shell mass as a function of radius as

$$M(r) = m(U)|_\Sigma\,, \quad M' = m'\left.\frac{\partial U}{\partial r}\right|_\Sigma = -\frac{2}{f}m', \tag{3.104}$$

equation (3.103) can be written as an ordinary differential equation of the first order for the on-shell mass M

$$\frac{v_0^2}{4r^2+v_0^2} - \frac{2}{f}M'(r) = 0. \tag{3.105}$$

The solution is obtained easily enough; it is

$$M(r) = \frac{1}{r}\left[c_1\sqrt{4r^2+v_0^2} - \frac{v_0^2}{8}\right], \tag{3.106}$$

where c_1 is a constant of integration, undetermined at the moment.

There is still one remaining condition to be satisfied

$$[\mathcal{K}_{\theta\theta}] = \frac{r}{2}f - \frac{2r^2+uv_0}{4r} = 0, \tag{3.107}$$

which requires the on-shell mass to be

$$M(r) = -\frac{uv_0}{4r}. \tag{3.108}$$

But recall that on the Roberts side, retarded time u is related to radius r and advanced time v by

$$u = \frac{1}{2}v - \frac{1}{2}\sqrt{4r^2+v^2}. \tag{3.109}$$

Using this expression, one can see that conditions (3.103) and (3.107) are both satisfied by taking the on-shell mass to be

$$M(r) = \frac{1}{r}\left[\frac{v_0}{8}\sqrt{4r^2+v_0^2} - \frac{v_0^2}{8}\right], \tag{3.110}$$

i.e. setting the constant of integration to $c_1 = v_0/8$.

Thus, the outgoing Vaidya solution can be smoothly matched to the Roberts solution across any null surface $v = \text{const}$. The global structure of the resulting spacetime is more complicated than for the Roberts solution alone, as is shown on Fig. 4. The null singularity disappears when the influx of the scalar field is switched off, so apparently the null singularity in the center of Roberts spacetime is sustained by the incoming flux of the scalar field.

Before we move on to another topic, there is still one more thing we would like to clarify. It is the question of how the null coordinate U in the outgoing Vaidya

solution is related to the null coordinate u in the Roberts solution. Integrating $f\,dU + 2dr = 0$ on the null shell Σ, we have

$$U = -\int \frac{2dr}{f(r)} = -2r - v_0 \ln\left[2r + \sqrt{4r^2 + v_0^2}\right] + \text{const.} \tag{3.111}$$

But $r = \sqrt{u^2 - uv_0}$ from the Roberts side, so we have the desired relationship

$$U = -2\sqrt{u^2 - uv_0} - v_0 \ln\left[\frac{2\sqrt{u^2 - uv_0} + v_0 - 2u}{v_0}\right], \tag{3.112}$$

where we fixed the constant so that $U = 0$ when $u = 0$.

3.8 Quantum corrections

In the course of critical collapse in the classical general relativity, arbitrarily small black holes and correspondingly large spacetime curvatures can be produced. In these regions of spacetime, quantum effects such as vacuum polarization and particle creation become important. It would be extremely interesting to study how the critical behavior is modified by the quantum theory of gravity. Unfortunately, a complete and self-consistent quantum gravity theory is beyond current techniques, but at least the first order quantum corrections to critical solutions should be examined.

There have been numerical simulations of several scalar field models incorporating semi-classical corrections [4, 16, 105], as well as analytical analysis of quantum effects in spherically symmetric self-similar collapse using dimensional reduction and the trace anomaly method [29, 20][1]. The general expectations of the effect of quantum corrections is that they will break self-similarity of the critical solution in the regions of strong curvature, and will prohibit arbitrarily small black holes, turning on black hole formation at a mass gap of the order of the Planck mass scale. These seem to be supported by the above studies, although the results are still not conclusive.

It is beyond the scope of this thesis to go into detailed calculations of quantum corrections to the scalar field model of critical behavior we considered. However, we are in position to estimate the particle production by a naked singularity in the self-similar collapse of a massless scalar field using the geometrical optics approximation of Ford and Parker [38]. For simplicity, we will calculate the particle production rates for some hypothetical massless scalar field, different than the background field ϕ of the Roberts solution.

[1] Quantum effects in related RST dilaton gravity model were studied in [114]. However, that model does not exhibit critical behavior classically.

The total power radiated across a sphere of radius r at late times is

$$P = \int \langle 0|T^r_t|0\rangle \, r^2 \sin\theta \, d\theta d\phi, \tag{3.113}$$

and contains contributions from different spherical harmonics. Considering only the s-mode contribution, and neglecting scattering on the potential barrier in the geometrical optics approximation, Ford and Parker argue that in asymptotically flat space time, the rate at which the massless scalar field particles are created is

$$P_0 = \frac{1}{24\pi}\left[\frac{F'''}{(F')^3} - \frac{3}{2}\frac{(F'')^2}{(F')^4}\right], \tag{3.114}$$

where function $U = F(v)$ describes the light ray propagating on the background from null ray $v = \text{const}$ at $\mathcal{J}^-$ to null ray $U = \text{const}$ at $\mathcal{J}^+$. Recall that the Roberts spacetime can be made asymptotically flat by gluing it between flat spacetime and outgoing Vaidya solution, as it was done in Section 3.7. We already obtained the relationship between null coordinates at $\mathcal{J}^-$ and $\mathcal{J}^+$ for radial light rays in the resulting spacetime; it is given by equation (3.112). In view of this, the s-mode quantum flux produced by the naked singularity in the Roberts solution is

$$P_0 = -\frac{1}{768\pi}\frac{(3v_0 - 8u)v_0}{(v_0 - u)^3 u}. \tag{3.115}$$

It is positive and diverges as u^{-1} as one approaches the singularity. By contrast, the outgoing scalar field flux in the Vaidya solution is regular. This simple calculation indicates that quantum particle production becomes overwhelmingly important near the singularity in our model, and quantum back-reaction might cause the singularity of the classical solution to disappear. This result is in agreement with conclusions of Refs. [29, 20]. In all likelihood, a similar effect would probably occur in any critical solution.

Having made this remark about the importance of quantum effects, we return to our main subject, the study of critical collapse of a massless scalar field in the classical gravity.

Chapter 4

Spherically Symmetric Perturbations

In this chapter, we study the spherically symmetric perturbations of the continuously self-similar scalar field solution we discussed in the last chapter. In order to analyze the stability of the Roberts solution, we derive the perturbation equations and determine the spectrum of the growing perturbation modes. Remarkably, the perturbation problem allows exact analytical treatment. The perturbation spectrum of the growing spherically symmetric modes on the Roberts background turns out to be not discrete, but occupying a continuous region of the complex plane.

We choose to perturb the full system of Einstein-scalar field equations

$$R_{\mu\nu} = 2\phi_{,\mu}\phi_{,\nu}, \quad \Box\phi = 0, \tag{4.1}$$

rather than spherically reduced equations (3.4) derived in the last chapter, so as not to change the treatment and notation drastically when we consider the non-spherical perturbations in the next chapter. To avoid possible gauge-related issues, we will work in fully gauge-invariant formalism.

The results of this chapter first appeared in [39], although here we mostly follow the notation and simpler derivation of [42].

4.1 Gauge-invariant linear perturbations

In this section, we outline how spherically symmetric perturbations of the Roberts solution (3.62) are described in gauge-invariant formalism. A general spherically-symmetric metric perturbation is

$$\delta g_{\mu\nu} dx^\mu dx^\nu = h_{uu}\, du^2 + 2h_{uv}\, du\, dv + h_{vv}\, dv^2 + r^2 K\, d\Omega^2, \tag{4.2}$$

while a general perturbation of the scalar field is

$$\delta\phi = \varphi. \tag{4.3}$$

Under a (spherically-symmetric) gauge transformation generated by the vector

$$\xi^\mu = (A, B, 0, 0), \tag{4.4}$$

the metric and scalar field perturbations transform as

$$\Delta g_{\mu\nu} = \pounds_\xi g_{\mu\nu}, \quad \Delta\phi = \pounds_\xi \phi. \tag{4.5}$$

The explicit expressions for the change in the perturbation amplitudes under the gauge transformation generated by the vector ξ are

$$\begin{aligned} \Delta h_{uu} &= -2B_{,u}, \\ \Delta h_{uv} &= -A_{,u} - B_{,v}, \\ \Delta h_{vv} &= -2A_{,v}, \\ r^2 \Delta K &= (2u - v)A - uB, \\ 2r^2 \Delta\varphi &= vA - uB. \end{aligned} \tag{4.6}$$

Out of four metric and one matter perturbation amplitudes one can build the total of three *gauge-invariant* quantities, one describing matter perturbations

$$f = \frac{K}{2} - \varphi + \frac{1}{2u}\int h_{vv}\, dv, \tag{4.7}$$

and the other two describing metric perturbations

$$\rho = (r^2 K)_{,uv} + h_{vv} - h_{uv} - u h_{uu,v}/2 + (2u - v) h_{vv,u}/2, \tag{4.8}$$

$$\sigma = h_{uv} - \frac{1}{2}\int h_{vv,u}\, dv - \frac{1}{2}\int h_{uu,v}\, du. \tag{4.9}$$

The linearized Einstein-scalar field equations can then be rewritten completely in terms of these gauge-invariant quantities. Once the gauge-invariant quantities are identified, one is free to switch between various gauges. We conclude this section by discussing two particularly convenient choices.

Field gauge ($K = h_{vv} = 0$):

The scalar field perturbation coincides with the gauge-invariant quantity f in this gauge, and expressions for other gauge-invariant quantities simplify considerably:

$$\begin{aligned} f &= -\varphi, \\ \rho &= -h_{uv} - u h_{uu,v}/2, \\ \sigma &= h_{uv} - \frac{1}{2}\int h_{uu,v}\, du. \end{aligned} \tag{4.10}$$

The linearized Einstein-scalar field equations are at their simplest in this gauge.

Null gauge ($h_{uu} = h_{vv} = 0$):

This gauge was used in the original analysis of spherically-symmetric perturbations of the Roberts solution [39]. The motivation behind this gauge choice is that coordinates u and v remain null in the perturbed spacetime. The expressions for gauge-invariant quantities are quite simple here as well:

$$\begin{aligned} f &= K/2 - \varphi, \\ \rho &= (r^2 K)_{,uv} - h_{uv}, \\ \sigma &= h_{uv}. \end{aligned} \tag{4.11}$$

4.2 Decoupling of perturbation equations

The simplest way to derive the perturbation equations for the Roberts background is to consider the linearized Einstein-scalar field equations

$$\delta R_{\mu\nu} = 4\phi_{(,\mu}\delta\phi_{,\nu)}, \quad \delta(\Box\phi) = 0 \tag{4.12}$$

in the field gauge ($K = h_{vv} = 0$). After carrying out the calculation of perturbed Ricci tensor components and keeping only linear terms, one finds that

$$uh_{uu,u} + (2u - v)h_{uu,v} - 2(2u - v)h_{uv,u} - 4v\varphi_{,u} = 0, \tag{4.13a}$$

$$-u(u - v)h_{uu,vv} + 2u(u - v)h_{uv,uv} + uh_{uu,v} - 2v\varphi_{,v} + 2u\varphi_{,u} = 0, \tag{4.13b}$$

$$h_{uv,v} + 2\varphi_{,v} = 0, \tag{4.13c}$$

$$2h_{uv} - uh_{uu,v} = 0, \tag{4.13d}$$

for uu, uv, vv, and $\theta\theta$ components of the Einstein equations correspondingly. The perturbation of the scalar field wave equation yields

$$4u(u - v)\varphi_{,uv} + 2(2u - v)\varphi_{,v} - 2u\varphi_{,u} - uh_{uu,v} = 0. \tag{4.13e}$$

The perturbation equations (4.13) inherit their structure from the general Einstein-scalar equations (4.1); equations (4.13c,4.13d,4.13e) are dynamical equations, while equation (4.13a) is a constraint. Equation (4.13b) is redundant, and will be ignored. The equation (4.13c) has a particularly simple form, and can be integrated immediately:

$$h_{uv} + 2\varphi = C, \tag{4.14}$$

where C is a constant of integration to be fixed by initial conditions.

Now let us restore the gauge-invariant variables in the equations using expressions (4.10) for the field gauge. Combining equations (4.13e) and (4.13d), and using the equation (4.14), one obtains the master differential equation for the scalar field perturbation

$$2u(u-v)f_{,uv} + (2u-v)f_{,v} - uf_{,u} - 2f = C. \tag{4.15}$$

From equation (4.13d) alone, and its combination with equations (4.14,4.13c) one finds the remaining two equations relating gauge-invariant metric perturbations to the gauge-invariant scalar field perturbation

$$\sigma_{,u} = 2f_{,u} + 2f/u + C, \quad \rho = 0. \tag{4.16}$$

If one wishes to attach a perturbed Roberts spacetime to the flat one, as we do, and still have no singularity at the junction, the null shell matching conditions written down in Section 3.7.1 will place boundary conditions on the perturbation values at the junction surface. One can show [42] that these boundary conditions on perturbations are the vanishing of gauge-invariant perturbation amplitudes on the $v = 0$ hypersurface

$$f = \rho = \sigma = 0 \text{ on } v = 0. \tag{4.17a}$$

(This condition determines the constant of integration C to be zero.) We also demand well-behavedness of perturbations at infinity for perturbation expansion to be valid

$$f, \rho, \sigma \text{ are bounded at } \mathcal{J}^{\pm}. \tag{4.17b}$$

Thus, the study of spherically symmetric perturbations of the Roberts solution is reduced to solving single partial differential equation

$$2u(u-v)f_{,uv} + (2u-v)f_{,v} - uf_{,u} - 2f = 0, \tag{4.18}$$

subject to boundary conditions (4.17). Once the solution is obtained, the gauge-invariant metric perturbations are found using equations (4.16). Alternatively, if we choose to work in the field gauge, the metric and scalar field perturbation amplitudes are trivial to obtain from gauge-invariant quantity f

$$\varphi = -f, \quad h_{uv} = 2f, \quad h_{uu,v} = -4f/u. \tag{4.19}$$

4.3 Separation of variables

The master equation for the scalar field perturbation (4.18) we derived in the previous section can be further simplified by exploiting self-similarity of the background Roberts solution. This symmetry of the background guarantees that the scale and spatial variables will separate. In other words, if we rewrite equation (4.18) in scaling coordinates defined in Section 3.6.3,

$$\mathcal{D}f(x,s) \equiv (1-e^{-2x})\frac{\partial^2 f}{\partial x^2} + 2\frac{\partial^2 f}{\partial x\,\partial s} + 2\frac{\partial f}{\partial s} - 4f = 0, \tag{4.20}$$

we find that the coefficients of the differential equation do not depend on the scale s, and so the problem can be reduced to one dimension by applying a formal Laplace transform with respect to the scale variable s. For the Laplace transform of f we have

$$F(x,k) = \int_0^\infty f(x,s)e^{-ks}ds, \tag{4.21}$$

with the inverse transformation being

$$f(x,s) = \frac{1}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty} F(x,k)e^{ks}dk. \tag{4.22}$$

The Laplace transform can be done provided that f can be bounded by an exponential function of s (that is, there exist constants M, κ_0 such that $|f(x,s)| \leq Me^{\kappa_0 s}$), which is a physically reasonable condition. The contour of integration in the complex k-plane for the inverse transform (4.22) must be taken somewhere to the right of κ_0 ($\kappa > \kappa_0 \geq 0$). The properties of functions of complex variables will guarantee that the result of integration is independent of the particular contour choice.

When applying the Laplace transform to the differential operator,

$$\mathcal{L}_s\left[\frac{\partial f}{\partial s}\right] = kF - f(s=0), \tag{4.23}$$

so the initial conditions of the original problem will enter as source terms on the right hand side. Therefore the Laplace transform of the equation (4.20) is

$$\mathcal{D}_k F(x,k) = h(x), \tag{4.24}$$

where $\mathcal{D}_k = \mathcal{L}_s\mathcal{D}$ is now an ordinary differential operator, algebraic in k, and h contains information about the initial shape of the wavepacket at $s = 0$. Boundary conditions on equation (4.24) are inherited from the original problem by Laplace transformation.

4.4 Growing modes

We will postpone the analysis of the evolution of arbitrary wavepackets on the Roberts background until Chapter 6, and will only look for the free growing perturbation modes satisfying both boundary conditions (4.17) here. They will determine the spectrum of the spherically symmetric perturbations of the Roberts solution.

The explicit form of the operator $\mathcal{D}_k$ is the simplest when expressed in a slightly different spatial coordinate, related to the old one by

$$y = e^{2x} = 1 - \frac{v}{u}. \tag{4.25}$$

In these coordinates, the differential operator $\mathcal{D}_k$ is hypergeometric in nature

$$\mathcal{D}_k F(y,k) \equiv y(1-y)\ddot{F} + [1-(k+1)y]\dot{F} - (k/2-1)F = h, \tag{4.26}$$

where here and later a dot denotes the derivative with respect to y ($\dot{} = d\,/dy$). The right hand side h depends on the initial conditions as

$$h(y) = -y\dot{f}(y, s=0) - \tfrac{1}{2} f(y, s=0). \tag{4.27}$$

For the initial conditions (4.17a), the right hand side h vanishes, and F satisfies the homogeneous hypergeometric equation (A.1) with coefficients

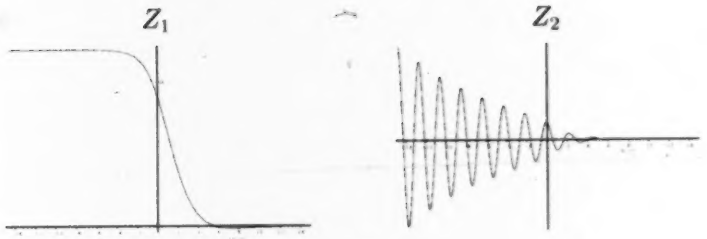
$$\begin{aligned} &c = 1, \quad a+b = k, \quad ab = k/2-1, \\ &a, b = 1/2\,(k \mp \sqrt{k^2-2k+4}). \end{aligned} \tag{4.28}$$

The properties of the hypergeometric equation are well-understood (see, for example, [8]), so equation (4.26) can be solved analytically and the perturbation spectrum determined exactly. We collect some facts about the hypergeometric equation and its solutions which are of immediate use to us in Appendix A.

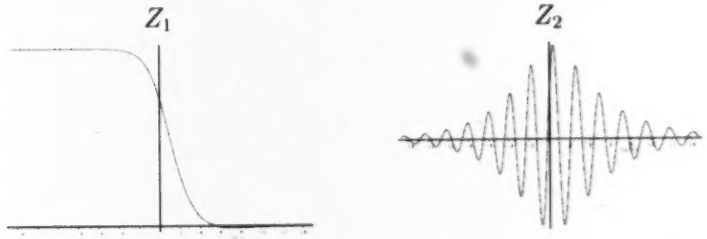
The general solution of the hypergeometric equation is a linear combination of any two solutions from the set (A.2). Imposing boundary conditions (4.17) on the general solution leads to the perturbation spectrum. When doing this, we will use asymptotics (A.4) to analyze the behavior of the solution in the boundary region[1]. The behavior of solutions Z_1 and Z_2 for several particular values of k is illustrated in Fig. 5.

First, let us examine boundary conditions at $y = 1$. Observe that at $y = 1$, Z_2 diverges as a power for $\mathrm{Re}\, k > 1$, logarithmically for $k = 1$, does not have a limit

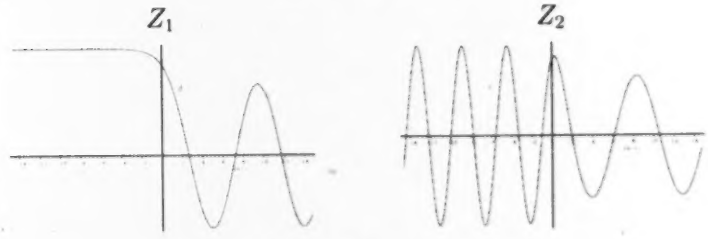
[1]The case when $k = 1$ is special. In this case the solution degenerates and Z_2 defined by (A.2) coincides with Z_1 identically. In order to deal with this situation we continue to use Z_1 as earlier, but denote by Z_2 a solution independent of Z_1. It is easy to verify that it is logarithmically divergent at $y = 1$.



(a) $k = \frac{11}{10} + 3i$



(b) $k = \frac{3}{4} + 3i$



(c) $k = 1 + i\sqrt{2}$

(d) $k = \frac{3}{4} + i\sqrt{2}$

Figure 5: Field perturbation profiles on a slice $u = \text{const}$: (a) for a typical value of k inside region A in Fig. 6, (b) for a typical value of k inside region F in Fig. 6, (c) for a value of k at the endpoint of region F in Fig. 6, (d) for a typical value of k inside region C in Fig. 6. The horizontal coordinate on the plots is $\ln v$.

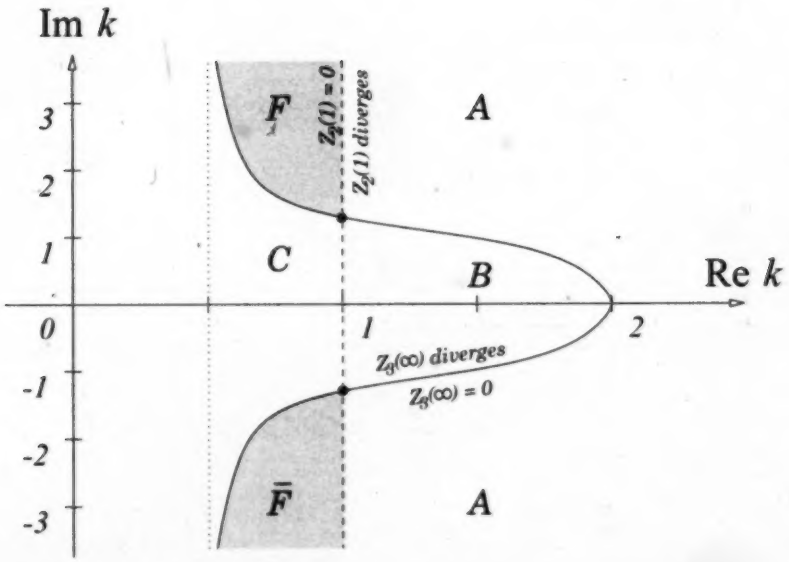


Figure 6: Complex perturbation spectrum. Values of k to the left of the solid line are prohibited by the boundary conditions at infinity, to the right of the broken line by the initial conditions at $y = 1$. Values in the region of intersection (the shaded regions F and $\bar{F}$) are allowed, and constitute the perturbation spectrum.

for $\mathrm{Re}\, k = 1$, $\mathrm{Im}\, k \neq 0$, and converges (to zero) only for $\mathrm{Re}\, k < 1$. The solutions that satisfy the initial condition $F(1) = 0$ are

$$F(y) \propto Z_2(y), \quad \mathrm{Re}\, k < 1. \tag{4.29}$$

From equation (A.5) we see that Z_2 (and Z_1) is connected to Z_3, Z_4 by a linear relation with non-zero coefficients, so the boundary conditions at infinity will only be satisfied if both Z_3, Z_4 do not blow up, i.e. if both $\mathrm{Re}\, a > 0$ and $\mathrm{Re}\, b > 0$. $\mathrm{Re}\, b$ is positive by construction, while the region $\mathrm{Re}\, a > 0$ of the complex k-plane is divided from the region $\mathrm{Re}\, a < 0$ by the curve $\mathrm{Re}\, a = 0$. Written in terms of the real and imaginary parts of k, it has the form

$$(\mathrm{Im}\, k)^2 = \frac{\mathrm{Re}\, k\,(2 - \mathrm{Re}\, k)}{1 - (2\mathrm{Re}\, k)^{-1}}, \tag{4.30}$$

and is shown in Fig. 6: to the left of the solid curve $\mathrm{Re}\, a < 0$, and to the right $\mathrm{Re}\, a > 0$.

Combining restrictions on k placed by boundary conditions at $y = 1$ and $y = \infty$ we see that to satisfy all boundary conditions (4.17), k must lie in the shaded region $F \cup \bar{F}$ of the complex plane, and for any $k \in F \cup \bar{F}$,

$$F(y) = (1 - y)^{1-k} \mathcal{F}(1 - a, 1 - b; 2 - k; 1 - y) \tag{4.31}$$

is a solution of equation (4.26) obeying boundary conditions (4.17).

Thus, the spectrum of the spherically symmetric perturbations of the Roberts solution is not discrete, but continuous, and occupies the region of the complex plane

$$\frac{1}{2} < \mathrm{Re}\, k < 1, \quad |\mathrm{Im}\, k| > \sqrt{\frac{\mathrm{Re}\, k\,(2 - \mathrm{Re}\, k)}{1 - (2\mathrm{Re}\, k)^{-1}}}, \tag{4.32}$$

as illustrated in Fig. 6. These growing modes are responsible for driving the perturbation away from the Roberts spacetime. We will return to this subject in Chapter 6, where we will investigate the mechanism of this departure in more detail.

Chapter 5

Beyond Spherical Symmetry

An important question is how generic the critical behavior is with respect to initial data, or, in phase space language, how big is the basin of attraction of the critical solution. So far most of the work on critical gravitational collapse, numeric or analytic, has been restricted to the case of spherical symmetry, simply because of the enormous difficulties in treating fully general non-symmetric solutions of Einstein equations. A natural concern is whether the critical phenomena observed so far are limited to spherical symmetry, and whether the evolution of non-spherical data will lead to the same results. The numerical study of Abrahams and Evans on axisymmetric gravitational wave collapse [1] and recent numerical perturbation calculations by Gundlach [61, 64] give numerical evidence for the claim that critical phenomena are not restricted to spherical symmetry, and that the critical solutions are indeed attractors in the full phase space.

In this chapter we consider fully general non-spherical perturbations of the Roberts solution in a gauge-invariant formalism. Due to the symmetries of the background, the linear perturbation equations decouple and the variables separate, so an exact analytical treatment is possible. We find that there are no growing perturbation modes apart from spherically symmetric ones described in the last chapter. So all the non-sphericity of the initial data decays in the collapse of the scalar field, and only the spherically symmetric part will play a role in the critical behavior.

The material presented in this chapter mostly follows [41].

5.1 Gauge-invariant linear perturbations

To avoid complicated gauge issues of fully general perturbations, we will use the gauge-invariant formalism developed by Gerlach and Sengupta [51, 52]. This formalism describes perturbations around a general spherically symmetric back-

ground

$$g_{\mu\nu}dx^\mu dx^\nu = \gamma_{AB}dx^A dx^B + r^2\omega_{ab}dx^a dx^b, \tag{5.1}$$

which in our case we take to be the Roberts solution (3.62). Here and later capital Latin indices take values $\{0,1\}$, and lower-case Latin indices run over angular coordinates. γ_{AB} and r are defined on a spacetime two-manifold, while ω_{ab} is the metric of the unit two-sphere.

Because the background spacetime is spherically symmetric, perturbations around it can be decomposed in spherical harmonics. Scalar spherical harmonics $Y_{lm}(\theta,\varphi)$ have even parity under spatial inversion, while vector spherical harmonics $S_{lm\,a}(\theta,\varphi) \equiv \epsilon_a{}^b Y_{lm,b}$ have odd parity. We will only concern ourselves with even-parity perturbations here, since odd-parity perturbations cannot couple to scalar field perturbations. We will focus on non-spherical perturbation modes ($l \geq 1$), as the spherically symmetric case ($l = 0$) was studied in the last chapter. For brevity, angular indices l, m and the summation over all harmonics will be suppressed from now on. The most general even-parity metric perturbation is

$$\begin{aligned} \delta g_{\mu\nu}dx^\mu dx^\nu &= h_{AB}Y dx^A dx^B \\ &\quad + h_A Y_{,b}(dx^A dx^b + dx^b dx^A) \\ &\quad + r^2[KY\omega_{ab} + GY_{:ab}]dx^a dx^b, \end{aligned} \tag{5.2}$$

and the scalar field perturbation is

$$\delta\phi = FY. \tag{5.3}$$

As you can see, metric perturbations are described by a two-tensor h_{AB}, a two-vector h_A, and two two-scalars K and G; the scalar field perturbation is described by a two-scalar F. However, these perturbation amplitudes do not have direct physical meaning, as they change under the (even-parity) gauge transformation induced by the infinitesimal vector field

$$\xi_\mu dx^\mu = \xi_A Y dx^A + \xi Y_{,a}dx^a. \tag{5.4}$$

One can construct two gauge-invariant quantities from the metric perturbations

$$\begin{aligned} k_{AB} &= h_{AB} - 2p_{(A|B)}, \\ k &= K - 2v^A p_A, \end{aligned} \tag{5.5}$$

and one from the scalar field perturbation

$$f = F + \phi^{,A}p_A, \tag{5.6}$$

where

$$v_A = \frac{r_{,A}}{r}, \quad p_A = h_A - \frac{r^2}{2} G_{,A}. \tag{5.7}$$

Only gauge-invariant quantities have physical meaning in the perturbation problem. All physics of the problem, including the equations of motion and boundary conditions, should be written in terms of these gauge-invariant quantities. Once gauge-invariant quantities have been identified, one is free to convert between gauge-invariant perturbation amplitudes and their values in whatever gauge choice one desires.

5.2 Perturbation equations

We will work in longitudinal gauge ($h_A = G = 0$), which is particularly convenient since perturbation amplitudes in it are just equal to the corresponding gauge-invariant quantities. The above condition fixes the gauge uniquely for non-spherical modes[1]. The perturbed metric in longitudinal gauge is

$$ds^2 = h_{uu} Y du^2 - 2(1 - h_{uv} Y) du\, dv + h_{vv} Y dv^2 + (1 + KY) r^2 d\Omega^2, \tag{5.8}$$

and the perturbed scalar field is

$$\phi = \frac{1}{2} \ln \left[1 - \frac{v}{u}\right] + FY. \tag{5.9}$$

Substituting these into Einstein-scalar field equations

$$E_{\mu\nu} \equiv R_{\mu\nu} - 2\phi_{,\mu}\phi_{,\nu} = 0, \quad \Box\phi = 0, \tag{5.10}$$

and keeping only linear terms, we obtain expressions for components of the perturbation equations calculated in the longitudinal gauge. The non-trivial components of perturbed Einstein-scalar field equations are given below:

$$\begin{aligned} E_{uu} = \frac{1}{2}\Big[& -2(u^2 - uv)K_{,uu} + u h_{uu,u} \\ & + (2u - v)(h_{uu,v} - 2h_{uv,u} - 2K_{,u}) \\ & - 4vF_{,u} + l(l+1)h_{uu}\Big] \frac{Y}{u^2 - uv}, \end{aligned} \tag{5.11a}$$

[1] There is some gauge freedom left over for the $l = 0$ mode, but in this chapter, we are only concerned with higher l modes.

$$E_{uv} = -\frac{1}{2}\Big[(u^2 - uv)(h_{uu,vv} - 2h_{uv,vu} + h_{vv,uu} + 2K_{,vu}) \tag{5.11b}$$
$$- uh_{uu,v} + (2u - v)(h_{vv,u} + K_{,v}) - uK_{,u}$$
$$+ 2vF_{,v} - 2uF_{,u} - l(l+1)h_{uv}\Big]\frac{Y}{u^2 - uv},$$

$$E_{vv} = \frac{1}{2}\Big[- 2(u^2 - uv)K_{,vv} + 2uh_{uv,v} \tag{5.11c}$$
$$- uh_{vv,u} - (2u - v)h_{vv,v} + 2uK_{,v}$$
$$+ 4uF_{,v} + l(l+1)h_{vv}\Big]\frac{Y}{u^2 - uv},$$

$$E_{u\theta} = -\frac{1}{2}\Big[(u^2 - uv)(h_{uu,v} - h_{uv,u} + K_{,u}) \tag{5.11d}$$
$$+ (2u - v)h_{uv} + 2vF\Big]\frac{Y_{,\theta}}{u^2 - uv},$$

$$E_{u\varphi} = -\frac{1}{2}\Big[(u^2 - uv)(h_{uu,v} - h_{uv,u} + K_{,u}) \tag{5.11e}$$
$$+ (2u - v)h_{uv} + 2vF\Big]\frac{Y_{,\varphi}}{u^2 - uv},$$

$$E_{v\theta} = -\frac{1}{2}\Big[(u - v)(-h_{uv,v} + h_{vv,u} + K_{,v}) - h_{uv} - 2F\Big]\frac{Y_{,\theta}}{u - v}, \tag{5.11f}$$

$$E_{v\varphi} = -\frac{1}{2}\Big[(u - v)(-h_{uv,v} + h_{vv,u} + K_{,v}) - h_{uv} - 2F\Big]\frac{Y_{,\varphi}}{u - v}, \tag{5.11g}$$

$$E_{\theta\theta} = \frac{1}{2}\Big[2(u^2 - uv)K_{,vu} - uh_{uu,v} + (2u - v)(h_{vv,u} + 2K_{,v}) \tag{5.11h}$$
$$- 2uK_{,u} - 2h_{uv} + 2h_{vv} - 2K + l(l+1)K\Big]Y + h_{uv}Y_{,\theta\theta},$$

$$E_{\varphi\varphi} = \sin^2\theta\, E_{\theta\theta}, \tag{5.11i}$$

$$E_{\theta\varphi} = h_{uv}(Y_{,\theta\varphi} - \cot\theta\, Y_{,\varphi}), \tag{5.11j}$$

$$\Box\phi = \frac{1}{2}\Big[- 4(u^2 - uv)F_{,vu} + uh_{uu,v} - vh_{vv,u} + uK_{,u} - vK_{,v} \tag{5.11k}$$
$$+ 2uF_{,u} - 2(2u - v)F_{,v} - 2l(l+1)F\Big]\frac{Y}{u^2 - uv}.$$

By inspection of the $\theta\varphi$ component (5.11j) of the perturbed Einstein-scalar field equations (5.11), it is clear that the equations of motion require that $h_{uv} = 0$ for $l \geq 1$. With the change of notation $h_{uu} = U$ and $h_{vv} = V$, the remaining equations of motion for non-spherical modes are

$$4(u^2 - uv)F_{,vu} - uU_{,v} - uK_{,u} + vK_{,v} + vV_{,u} - 2uF_{,u} + 2(2u - v)F_{,v} + 2l(l+1)F = 0, \tag{5.12a}$$

$$-2(u^2 - uv)K_{,uu} + uU_{,u} + (2u - v)(U_{,v} - 2K_{,u}) - 4vF_{,u} + l(l+1)U = 0, \tag{5.12b}$$

$$-(u^2 - uv)(U_{,vv} + 2K_{,vu} + V_{,uu}) + uU_{,v} + uK_{,u} - (2u - v)(K_{,v} + V_{,u}) + 2uF_{,u} - 2vF_{,v} = 0, \tag{5.12c}$$

$$-2(u^2 - uv)K_{,vv} + 2uK_{,v} - uV_{,u} - (2u - v)V_{,v} + 4uF_{,v} + l(l+1)V = 0, \tag{5.12d}$$

$$2(u^2 - uv)K_{,vu} - uU_{,v} - 2uK_{,u} + (2u - v)(2K_{,v} + V_{,u}) - 2K + l(l+1)K + 2V = 0, \tag{5.12e}$$

$$(u^2 - uv)(U_{,v} + K_{,u}) + 2vF = 0, \tag{5.12f}$$

$$(u - v)(V_{,u} + K_{,v}) - 2F = 0. \tag{5.12g}$$

Equation (5.12a) comes from the scalar wave equation, and equations (5.12b–5.12g) are the uu, uv, vv, $\theta\theta$, $u\theta$, and $v\theta$ components of the Einstein equations, correspondingly. As usual with a scalar field, the system (5.12) has one redundant equation, so equation (5.12c) is satisfied automatically by virtue of other equations. Equations (5.12f) and (5.12g) are constraints, and the remaining four equations are dynamic equations for four perturbation amplitudes U, V, K, and F.

Boundary conditions for the system (5.12) are specified at $v = 0$ and at spatial infinity. Continuity of matching with flat spacetime at the hypersurface $v = 0$ requires the vanishing of the perturbations there. We also require well-behavedness of the perturbations at $\mathcal{J}^{\pm}$, so that the perturbation expansion holds. Thus, the boundary conditions are

$$\begin{aligned} &U = V = K = F = 0 \text{ at } v = 0, \\ &U, V, K, F \text{ are bounded at } \mathcal{J}^{\pm}. \end{aligned} \tag{5.13}$$

Equations (5.12) together with boundary conditions (5.13) constitute our eigenvalue problem.

5.3 Decoupling of perturbation equations

It is possible to decouple the dynamic equations (5.12a–5.12e) by combining them with the constraints (5.12f) and (5.12g), and their first derivatives. After somewhat cumbersome algebraic manipulations, which we will not show here, the system of linear perturbation equations (5.12) can be rewritten as

$$2(u^2 - uv)F_{,vu} - uF_{,u} + (2u - v)F_{,v} + \frac{2vF}{u - v} + l(l+1)F = 0, \tag{5.14a}$$

$$2(u^2 - uv)U_{,vu} + uU_{,u} + 3(2u - v)U_{,v} + l(l+1)U = 0, \tag{5.14b}$$

$$2(u^2 - uv)V_{,vu} - 3uV_{,u} - (2u - v)V_{,v} + l(l+1)V = 0, \tag{5.14c}$$

$$2(u^2 - uv)K_{,vu} - uK_{,u} + (2u - v)K_{,v} - 2K + l(l+1)K = -2V - \frac{4uF}{u - v}, \tag{5.14d}$$

$$uU_{,v} + uK_{,u} + \frac{2vF}{u - v} = 0, \tag{5.14e}$$

$$V_{,u} + K_{,v} - \frac{2F}{u - v} = 0. \tag{5.14f}$$

This decoupled system of partial differential equations can be further simplified by exploiting the continuous self-similarity of the background to separate spatial and scale variables, just as we did in the last chapter. With this intent, we rewrite equations (5.14) in terms of the scaling coordinates (3.71)

$$\frac{1}{2}(1 - e^{-2x})F_{,xx} + F_{,xs} + F_{,s} - 2(1 - e^{-2x})F + l(l+1)F = 0, \tag{5.15a}$$

$$\frac{1}{2}(1 - e^{-2x})U_{,xx} + U_{,xs} - 2U_{,x} - U_{,s} + l(l+1)U = 0, \tag{5.15b}$$

$$\frac{1}{2}(1 - e^{-2x})V_{,xx} + V_{,xs} + 2V_{,x} + 3V_{,s} + l(l+1)V = 0, \tag{5.15c}$$

$$\frac{1}{2}(1 - e^{-2x})K_{,xx} + K_{,xs} + K_{,s} - 2K + l(l+1)K = -2V - 4e^{-2x}F, \tag{5.15d}$$

$$U_{,x} - (1 - e^{2x})K_{,x} + 2e^{2x}K_{,s} - 4(1 - e^{2x})F = 0, \tag{5.15e}$$

$$K_{,x} - (1 - e^{2x})V_{,x} + 2e^{2x}V_{,s} + 4F = 0. \tag{5.15f}$$

We decompose the perturbation amplitudes into modes that grow exponentially with the scale s (which amounts to doing a Laplace transform on them) at the rate

$e^{\kappa s}$. For brevity, the perturbation mode subscript κ and the explicit summation over all modes is suppressed, so henceforth F, U, V, and K will refer to the amplitudes of the mode with eigenvalue κ.

The Laplace transformation converts the system of partial differential equations (5.15) into a system of ordinary differential equations, which is much easier to analyze:

$$\frac{1}{2}(1-e^{-2x})F'' + \kappa F' + \kappa F - 2(1-e^{-2x})F + l(l+1)F = 0, \tag{5.16a}$$

$$\frac{1}{2}(1-e^{-2x})U'' + (\kappa-2)U' - \kappa U + l(l+1)U = 0, \tag{5.16b}$$

$$\frac{1}{2}(1-e^{-2x})V'' + (\kappa+2)V' + 3\kappa V + l(l+1)V = 0, \tag{5.16c}$$

$$\frac{1}{2}(1-e^{-2x})K'' + \kappa K' + (\kappa-2)K + l(l+1)K = -2V - 4e^{-2x}F, \tag{5.16d}$$

$$U' - (1-e^{2x})K' + 2\kappa e^{2x}K - 4(1-e^{2x})F = 0, \tag{5.16e}$$

$$K' - (1-e^{2x})V' + 2\kappa e^{2x}V + 4F = 0. \tag{5.16f}$$

The prime denotes a derivative with respect to spatial variable x. By a change of variable (4.25), these equations can be cast into standard algebraic form, so that the system (5.16) becomes

$$y(1-y)\ddot{\Phi} + [3-(\kappa+3)y]\dot{\Phi} - [3\kappa/2 + l(l+1)/2]\Phi = 0, \tag{5.17a}$$

$$y(1-y)\ddot{U} + [1-(\kappa-1)y]\dot{U} - [-\kappa/2 + l(l+1)/2]U = 0, \tag{5.17b}$$

$$y(1-y)\ddot{V} + [1-(\kappa+3)y]\dot{V} - [3\kappa/2 + l(l+1)/2]V = 0, \tag{5.17c}$$

$$y(1-y)\ddot{K} + [1-(\kappa+1)y]\dot{K} - [\kappa/2 - 1 + l(l+1)/2]K = 2\Phi + V, \tag{5.17d}$$

$$\dot{U} + (y-1)\dot{K} + \kappa K + 2y\Phi - 2\Phi = 0, \tag{5.17e}$$

$$\dot{K} + (y-1)\dot{V} + \kappa V + 2\Phi = 0. \tag{5.17f}$$

The dot denotes a derivative with respect to y, and we redefined the scalar field perturbation amplitude as $F = y\Phi$ to cast the equations into standard table form. The boundary conditions (5.13) are

$$\begin{aligned} &U = V = K = \Phi = 0 \text{ at } y = 1, \\ &U, V, K, y\Phi \text{ are bounded at } y = +\infty. \end{aligned} \tag{5.18}$$

Imposed on the decoupled system of ordinary differential equations (5.17), these boundary conditions give an eigenvalue problem for the perturbation spectrum κ.

5.4 Perturbation spectrum

In the previous section we formulated an eigenvalue problem for the spectrum of non-spherical perturbations of the critical Roberts solution. We now proceed to solve it. Observe that equations (5.17a–5.17d) governing the dynamics of the perturbations are hypergeometric equations of the form (A.1). Equation (5.17d) is not homogeneous, but we will deal with that shortly. The hypergeometric equation coefficients are different for equations describing the perturbations Φ, U, V, and K; they are summarized in the table below.

	c	$a+b$	ab
Φ	3	$\kappa+2$	$\frac{3}{2}\kappa+\frac{1}{2}l(l+1)$
U	1	$\kappa-2$	$-\frac{1}{2}\kappa+\frac{1}{2}l(l+1)$
V	1	$\kappa+2$	$\frac{3}{2}\kappa+\frac{1}{2}l(l+1)$
K	1	κ	$\frac{1}{2}\kappa+\frac{1}{2}l(l+1)$

(5.19)

Once again, we refer the reader to Appendix A for the summary of properties of the hypergeometric equation and its solutions.

As we said before, imposing the boundary conditions (5.18) on solutions of equations (5.17) leads to a perturbation spectrum. We will now investigate what restrictions the boundary conditions place on the hypergeometric equation coefficients. The vanishing of perturbation amplitudes at $y = 1$ rules out Z_1 as a component of the solution and requires that $\mathrm{Re}\,(c-a-b) > 0$ to make Z_2 go to zero. The solution Z_2 has non-zero content of both Z_3 and Z_4 by virtue of (A.5), hence for it to be bounded at infinity, both Re a and Re b must be positive to guarantee convergence of Z_3 and Z_4. So, unless there is degeneracy, the boundary conditions translate to the following conditions on the hypergeometric equation coefficients:

$$\mathrm{Re}\,(c-a-b) > 0, \tag{5.20a}$$

$$\mathrm{Re}\,a, \mathrm{Re}\,b > 0. \tag{5.20b}$$

We are now ready to take on system (5.17). Take a look at equation (5.17c) for V. Condition (5.20a) for it is Re $\kappa < -1$, i.e. there are no growing V modes! With the amplitude of relevant V perturbation modes being zero, the constraints (5.17e) and (5.17f) become

$$K = -\frac{\dot{U}}{\kappa}, \quad \Phi = \frac{\ddot{U}}{2\kappa}, \tag{5.21}$$

and right hand side of equation (5.17d) can be absorbed by the left hand side, making the equation for K homogeneous (with $c = 2$). Indeed equations (5.17d) and (5.17a) for K and Φ are just derivatives of equation (5.17b) for U

$$y(1-y)\ddot{U} + [1 - (\kappa - 1)y]\dot{U} - [-\kappa/2 + l(l+1)/2]U = 0, \tag{5.22}$$

which is the homogeneous hypergeometric equation with coefficients

$$c = 1, \quad a + b = \kappa - 2, \quad ab = -\frac{1}{2}\kappa + \frac{1}{2}l(l+1). \tag{5.23}$$

Imposing the boundary condition at $y = 1$ for the solution of the above equation and its derivatives, which behave like

$$\left.\begin{array}{l} U \propto (1-y)^{3-\kappa} \\ K \propto (1-y)^{2-\kappa} \\ \Phi \propto (1-y)^{1-\kappa} \end{array}\right\} \text{ near } y = 1, \tag{5.24}$$

produces restriction on the non-spherical mode eigenvalue

$$\text{Re } \kappa < 1, \tag{5.25}$$

which is the strongest of the restrictions (5.20a) for equations for U, K, and Φ. But then

$$\text{Re } a + \text{Re } b = \text{Re } \kappa - 2 < -1, \tag{5.26}$$

and hence Re a and Re b can not be both positive, and so the boundary condition at infinity can not be satisfied. A more careful investigation of degenerate cases of relation (A.5) shows that the contradiction between boundary conditions at $y = 1$ and infinity still persists if $V = 0$. It can only be resolved by the trivial solution $U = K = \Phi = 0$. Thus we have shown that there are no growing non-spherical perturbation modes around the critical Roberts solution.

In fact, an even stronger statement is true. The contradiction between boundary conditions at $y = 1$ and infinity can not be resolved by a non-trivial solution so long as $V = 0$, i.e. so long as Re $\kappa \geq -1$. Hence non-trivial non-spherical perturbation modes of the critical Roberts solution must decay faster than e^{-s}.

Chapter 6

Continuous Self-Similarity Breaking

As we have mentioned, the critical solution often has additional symmetry besides the spherical symmetry, namely continuous or discrete self-similarity. This symmetry essentially amounts to the solution being independent of (in case of continuous self-similarity) or periodic in (in case of discrete self-similarity) one of the coordinates, a scale. The role of this symmetry in critical collapse is not particularly well understood.

In this chapter, we attempt to shed some light on the subject by investigating the formation of discretely self-similar structure in the gravitational collapse of a minimally coupled massless scalar field. Using linear perturbation analysis and Green's function techniques, we study evolution of the spherically symmetric scalar field configurations close to the continuously self-similar solution. Approximating late-time evolution via the method of stationary phase, we find that a generic growing perturbation departs from the Roberts solution in a universal way. In the course of the evolution, the initial continuous self-similarity of the background is broken into discrete self-similarity by the growing perturbation mode, reproducing the symmetries of the Choptuik solution. We are able to calculate the echoing period of the formed discretely self-similar structure on the Roberts background analytically, and its value is close to the result of numerical simulations for the Choptuik solution.

The material presented in this chapter follows [42].

6.1 Wave propagation on the Roberts background

We wish to study a scalar field wave propagating on the Roberts background. As we have shown in Chapter 4, the evolution of the spherically symmetric scalar

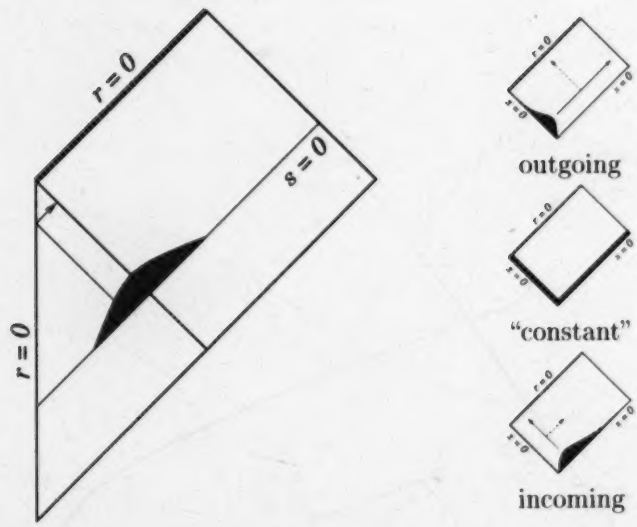


Figure 7: Wave propagation on the Roberts background: Initial conditions can be equivalently specified on the surface $s = 0$ extending to the center of the flat part of spacetime ($r = 0$), or on the $(x = 0) \cup (s = 0)$ wedge. By linearity, the wavepacket can be decomposed into three modes: outgoing, "constant", and incoming.

field perturbation is described by a single master equation (4.15) for the gauge-invariant field amplitude f. The wavepacket is specified by initial data either on some spacelike Cauchy surface or on an initial null surface.

Our choice for initial surface is $u = \text{const}$ ($s = 0$), which forms a complete null surface if extended to the center of the flat spacetime part, as illustrated in Fig. 7. The part of a pulse propagating through flat background evolves trivially, and can be equivalently replaced by specifying field values on the $v = 0$ hypersurface. Therefore, in our problem, the initial conditions for the linearized Einstein-scalar field equations are given on the $s = 0$ surface, while the boundary conditions are determined by junction conditions across the null shell $v = 0$, and the requirement that the perturbations be bounded at future infinity.

Specifics of the boundary conditions placed on the wave equation depend on the physical problem being considered. If the flat spacetime part $v < 0$ were unperturbed, the null shell junction conditions, discussed in Section 3.7.1, would require that $f = 0$ on the surface $v = 0$. If some part of the pulse propagates in the flat sector $v < 0$, f should be continuous across the surface $v = 0$. Essentially, we can specify the value of f on the wedge $(v = 0) \cup (s = 0)$, keeping in mind that the perturbation value should be bounded at future infinity. It is practical

to split the wavepacket into three components, as shown in Fig. 7, and consider outgoing, "constant" and incoming packets separately.

Recall that the scalar wave equation (4.15) can be reduced to the inhomogeneous hypergeometric equation (4.26) by a Laplace transformation, as it was done in Section 4.3. The initial conditions at $s = 0$ enter on the right hand side of (4.27), while the boundary conditions at $v = 0$ are inherited from the two-dimensional problem by a Laplace transformation with respect to s. We already discussed the existence of free growing modes – the solutions to wave equations with trivial boundary conditions – in Section 4.4. In this section, we derive the formal solution for evolution of an arbitrary wavepacket.

6.1.1 Outgoing wavepacket

The outgoing wavepacket is characterized by

$$f(x = 0, s) = f_O(s), \quad f(x, s = 0) = 0 \tag{6.1}$$

and is propagating outwards to future infinity, except for backscatter on the background curvature which goes towards the singularity at $s = +\infty$. The boundary conditions for equation (4.26) are inherited from the original problem by Laplace transformation $F(y, k) = \mathcal{L}_s f(y, s)$, and the right-hand side $h(y)$ is given by equation (4.27). For the outgoing wavepacket, they are

$$F(y = 1, k) = F_O(k), \quad h(y) = 0. \tag{6.2}$$

The general solution of the homogeneous form of equation (4.26) is

$$F(y, k) = A(k)Z_1(y, k) + B(k)Z_2(y, k), \tag{6.3}$$

where Z_1 and Z_2 are two linearly independent solutions of the homogeneous hypergeometric equation, in notation of Appendix A. To satisfy boundary conditions at $y = 1$, parameters A and B must be

$$A(k) = F_O(k), \quad B(k) = 0. \tag{6.4}$$

Therefore, the outgoing wavepacket solution is given by

$$f(y, s) = \frac{1}{2\pi i} \int_{\kappa - i\infty}^{\kappa + i\infty} F_O(k) \mathcal{F}(a, b; k; 1 - y) e^{ks}\, dk, \tag{6.5}$$

where $\mathcal{F}$ is the hypergeometric function.

If $f_O(s)$ does not grow exponentially as $s \to +\infty$ by itself, i.e. the image $F_O(k)$ does not have poles in Re $k > 0$ half-plane, than neither does $f(y, s)$. The outgoing wavepacket just propagates harmlessly out to future infinity, never growing enough to cause significant deviation of the solution from the Roberts background.

6.1.2 "Constant" Wavepacket

Even more trivial is the case of the "constant" wavepacket, characterized by

$$f(x=0,s)=C=f(x,s=0). \tag{6.6}$$

The boundary conditions and the initial term for equation (4.26) are

$$F(y=1,k)=C/k, \quad h(y)=-C/2. \tag{6.7}$$

The general solution of equation (4.26) with these boundary conditions is

$$F(y,k)=A(k)Z_1(y,k)+B(k)Z_2(y,k)+\frac{C}{k-2}, \tag{6.8}$$

and the boundary conditions at $y=1$ require that

$$A(k)=-\frac{2C}{k(k-2)}, \quad B(k)=0. \tag{6.9}$$

So the constant wavepacket solution is given by

$$f(y,s)=\frac{C}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty}\left[1-\frac{2}{k}\mathcal{F}(a,b;k;1-y)\right]\frac{e^{ks}}{k-2}\,dk, \tag{6.10}$$

and it does not grow as $s\to+\infty$ either.

6.1.3 Incoming Wavepacket

By far, the most physically interesting case is the incoming wavepacket characterized by

$$f(x=0,s)=0, \quad f(x,s=0)=f_I(x). \tag{6.11}$$

It propagates directly towards the singularity and is responsible for near-critical behavior and breaking of the solution away from the Roberts background, as we shall demonstrate. The boundary conditions and the initial term for equation (4.26) are

$$F(x=0,k)=0, \quad h(y)=-y\dot{f}_I(y)-f_I(y)/2. \tag{6.12}$$

To solve the inhomogeneous hypergeometric equation (4.26),

$$y(1-y)\ddot{F}+[1-(k+1)y]\dot{F}-(k/2-1)F=h, \tag{6.13}$$

with the boundary conditions

$$F(y=1,k)=0, \quad F(y=\infty,k) \text{ bounded}, \tag{6.14}$$

we must construct a Green's function out of the fundamental system of solutions of the homogeneous equation

$$\begin{aligned} Z_1(y) &= \mathcal{F}(a,b;k;1-y) \\ Z_2(y) &= (1-y)^{1-k}\mathcal{F}(1-a,1-b;2-k;1-y), \end{aligned} \tag{6.15}$$

where parameters a and b of hypergeometric equation depend on k as given by equation (4.28). The Wronskian of the above system is

$$W(y) = (k-1)y^{-1}(1-y)^{-k}, \tag{6.16}$$

and the Green's function is constructed as

$$\begin{aligned} G(y,\eta) &= AZ_1(y) + BZ_2(y) \\ &\qquad \pm \frac{1}{2p_0(\eta)W(\eta)}\left[Z_1(y)Z_2(\eta) - Z_2(y)Z_1(\eta)\right] \\ &= AZ_1(y) + BZ_2(y) \\ &\qquad \pm \frac{(1-\eta)^{k-1}}{2(k-1)}\left[Z_1(y)Z_2(\eta) - Z_2(y)Z_1(\eta)\right], \end{aligned} \tag{6.17}$$

where the coefficients A and B are to be determined by applying the boundary conditions, and the plus-minus sign is taken depending on the arguments of the Green's function

$$\pm = \begin{cases} +, & 1 \le y \le \eta \\ -, & \eta \le y < \infty \end{cases}. \tag{6.18}$$

The Green's function G satisfies $\mathcal{D}_k G(y,\eta) = \delta(y-\eta)$, and hence can be used to construct the solution of the inhomogeneous equation

$$F(y,k) = \int_1^\infty G(y,\eta)h(\eta)\,d\eta \tag{6.19}$$

As the initial value problem (6.13) is not self-adjoint, the Green's function (6.17) need not be symmetric in its arguments y and η. Note that the Green's function is calculated for a particular k-mode, and so depends on k, but we omitted the third argument in $G(y,\eta;k)$ for brevity.

We now proceed to apply boundary conditions to the Green's function (6.17), starting with the boundary conditions at $y=1$. The fundamental solution Z_1 goes to one there, while the behavior of Z_2 is fundamentally different depending

on the sign of Re $(1-k)$. If Re $k<1$, the real part of the exponent of $(1-y)$ in (6.15) is positive, and Z_2 goes to zero when $y=1$. If Re $k>1$, the real part of power of $(1-y)$ is negative, and hence Z_2 diverges when $y=1$. Substituting this into the Green's function, we get

$$\begin{aligned} G(y=1,\eta) &= \left[A+\frac{(1-\eta)^{k-1}}{2(k-1)}Z_2(\eta)\right]+ \\ &\quad +\left[B-\frac{(1-\eta)^{k-1}}{2(k-1)}Z_1(\eta)\right]\left\{\begin{array}{ll} 0, & \text{Re } k<1 \\ \infty, & \text{Re } k>1 \end{array}\right\}. \end{aligned} \tag{6.20}$$

The boundary conditions (6.14) require that $G(y=1,\eta)=0$, which uniquely fixes the coefficients

$$A=-\frac{(1-\eta)^{k-1}}{2(k-1)}Z_2(\eta), \quad B=\frac{(1-\eta)^{k-1}}{2(k-1)}Z_1(\eta), \tag{6.21}$$

provided Re $k>1$, which is precisely the region of the complex k-plane the contour of the inverse Laplace transformation should be in. (If Re $k<1$, coefficient B can be arbitrary.) With these coefficients, the Green's function (6.17) becomes

$$G(y,\eta)=\left\{\begin{array}{ll} 0, & y\le\eta \\ \frac{(1-\eta)^{k-1}}{k-1}\left[Z_1(\eta)Z_2(y)-Z_2(\eta)Z_1(y)\right], & \eta\le y \end{array}\right. . \tag{6.22}$$

The causality of wave propagation is apparent here: the wave at y is only influenced by the data from the past $\eta\le y$.

We see that the boundary conditions at $y=1$ already fix the Green's function (6.22), but we still have to satisfy the boundary conditions at infinity! One can show that they are satisfied automatically if (and only if) Re $a\ge 0$. Curves Re $k=1$ and Re $a=0$ split the complex k plane into several regions, as shown in Fig. 6 in Chapter 4. The Green's function (6.22) as written above is defined in region A, but could be analytically continued to the whole complex space. The obstructions the Green's function encounters on the boundaries between A and B and A and F are not poles, indeed, they are not even singular for regular (y,η). They are rather caused by the fact that the Green's function (6.22) fails to be applicable once you cross these boundaries; in the region B boundary conditions at infinity fail to be satisfied, and in the region F free modes (solutions of homogeneous equation, that is) exist that satisfy all boundary conditions, making the Green's function not unique. The existence of free modes in region F is at the heart of the matter, as they grow and will determine what will happen to the wavepacket at late times. (It is also possible to construct Green's function in region C, but since it has no bearing on our analysis, we will not do it here.)

Once the Green's function has been determined, it is simple to construct later-time evolution of the wavepacket from the initial data using equation (6.19)

$$F(y,k) = \int_1^y \frac{(1-\eta)^{k-1}}{k-1} \left[Z_1(\eta) Z_2(y) - Z_2(\eta) Z_1(y) \right] h(\eta)\, d\eta. \tag{6.23}$$

To get back from the complex k-plane dependence to the physical time evolution, one performs the inverse Laplace transformation

$$f(y,s) = \frac{1}{2\pi i} \int_{\kappa-i\infty}^{\kappa+i\infty} F(y,k) e^{ks} dk. \tag{6.24}$$

We emphasize again that a particular choice of the contour of integration is not important, as long as it is to the right of the obstructions on the complex plane, in our case region F. In practice, one chooses the contour so that the integral (6.24) is easier to evaluate. For some approximation to work, the contour should touch the obstruction, which means pushing it leftwards to the very edge of region F at Re $k = 1$.

6.2 Late-time behavior of incoming wavepacket

While expressions for f written down in the previous section formally solve the problem of wave propagation on the Roberts background, they are too complicated to be of practical use. In this section, we use the method of stationary phase to obtain the late-time (large s) asymptotic behavior for $f(y,s)$ and analyze several physically important regimes of the wavepacket evolution.

The method of the stationary phase deals with the approximate evaluation of Fourier-type integrals

$$f(\lambda) = \int_\alpha^\beta F(k) \exp[i\lambda S(k)]\, dk \tag{6.25}$$

for large positive parameter λ. It is based on the simple idea that where $\exp[i\lambda S(k)]$ is oscillating extremely rapidly and $F(k)$ is smooth, the oscillations will cancel out, and the only contributions to the integral will be from stationary points of phase $S(k)$, singular points of $F(k)$ and $S(k)$, and possibly end points.

Inverse Laplace transform integrals (6.24) are precisely of the above type, with phase $S(k) = k$ and the large parameter λ being the scale coordinate s. So the stationary phase method tells us that the asymptotic $s \to \infty$ behavior

of the solution $f(y,s)$ is given by singular points of $F(y,k)$ as a function of k, i.e. singular points of $G(y,\eta;k)$. Therefore the study of analytic properties of the Green's function (6.22) plays a key role in understanding the late-time evolution of the wavepacket. The possible sources of non-analyticity in Green's function are listed below:

- Branch points $k = 1 \pm i\sqrt{3}$ of $a, b = 1/2(k \pm \sqrt{k^2 - 2k + 4})$
- Poles at $k = 2 + n$ in Z_2 and $k = -n$ in Z_1 coming from the hypergeometric function
- The pole at $k = 1$ from the prefactor
- Power-law singularity of the type $(1-y)^{k-1}$

The problem with branches of the coefficients a, b is absent in the Green's function G because they only enter it through the first and second arguments of the hypergeometric function, and the hypergeometric series are written in terms of $ab = k/2 - 1$ and $a + b = k$ only. Various poles at integer values of k are all canceled out because of the antisymmetric way hypergeometric functions enter G. In fact, the only source of non-analyticity in G is the power-law singularity, and then only at $y, \eta \to 1$. Despite the appearance, $G(y,\eta;k)$ is an entire analytic function of k provided that y, η are regular points.

Since the small η region is important for the late-time evolution of the wavepacket, it is instructive to take a closer look at the approximation to the Green's function (6.22) there. Using the asymptotic behavior of Z_1, Z_2 near $\eta = 1$, given in Appendix A, we obtain

$$\begin{aligned} G(y,\eta\to 1) &\approx \frac{(1-\eta)^{k-1}}{k-1}\left[Z_2(y) - (1-\eta)^{1-k} Z_1(y)\right] \\ &= \frac{1}{k-1}\left[\left(\frac{1-\eta}{1-y}\right)^{k-1} \mathcal{F}_2(y) - \mathcal{F}_1(y)\right], \end{aligned} \tag{6.26}$$

where we introduce the short-hand notation

$$\begin{aligned} \mathcal{F}_1(y) &= \mathcal{F}(a,b;k;1-y), \\ \mathcal{F}_2(y) &= \mathcal{F}(1-a,1-b;2-k;1-y). \end{aligned} \tag{6.27}$$

Note that even though poles in the hypergeometric functions no longer cancel in (6.26), they are purely artifacts of the approximation, and should be ignored.

Now we will use the above approximation (6.26) for the Green's function to study the late-time evolution of the packet in several important regimes.

6.2.1 Evolution near $v = 0$

Let us first consider the behavior of the wavepacket near $y = 1$, that is, near the initial null surface $v = 0$. We take the asymptotic behavior of the initial term h to be the fairly generic power law

$$h(\eta) \propto (1-\eta)^\alpha, \tag{6.28}$$

which covers the usual case of functions analytic at $y = 1$ (via Taylor expansion), as well as the case of functions with a power-law singularity at $y = 1$, such as free modes Z_2. Then the approximation to the Green's function (6.26) gives the solution in the desired region,

$$F(y,k) \approx \int_1^y \frac{1}{1-k}\left[\left(\frac{1-\eta}{1-y}\right)^{k-1}\mathcal{F}_2(y) - \mathcal{F}_1(y)\right]\left(\frac{1-\eta}{1-y}\right)^\alpha (1-y)^\alpha\, d\eta. \tag{6.29}$$

The above integral can be explicitly evaluated using the change of variable

$$\zeta = \ln\left(\frac{1-\eta}{1-y}\right), \tag{6.30}$$

which leads to the answer

$$F(y,k) \approx -\int_{-\infty}^0 \frac{1}{1-k}\left[e^{(k-1)\zeta}\mathcal{F}_2(y) - \mathcal{F}_1(y)\right] e^{(1+\alpha)\zeta}(1-y)^{1+\alpha}\, d\zeta \tag{6.31}$$

$$= -(1-y)^{1+\alpha}\frac{1}{k-1}\left[\frac{\mathcal{F}_2(y)}{k+\alpha} - \frac{\mathcal{F}_1(y)}{1+\alpha}\right]. \tag{6.32}$$

To obtain the late-time evolution of the wavepacket near $y = 1$, we use the method of stationary phase. Observe that the only real singularity of $F(y,k)$ is the simple pole at $k = -\alpha$, with the rest being artifacts of the approximation (6.26). Therefore, the late-time behavior of the wavepacket is exponential in the scale s; indeed, it is given by

$$f(y,s) \approx \frac{(1-y)^{1+\alpha}}{1+\alpha}\mathcal{F}_2(y;-\alpha)e^{-\alpha s} = \frac{1}{1+\alpha}Z_2(y;-\alpha)\,e^{-\alpha s}. \tag{6.33}$$

There are several things worth noting about this result. First, it gives the correct answer for the evolution of the free mode. Since $Z_2(y;p) \approx (1-y)^{1-p}$, the initial term works out to be $h(y) \approx (1-p)(1-y)^{-p}$, so the above formula gives $f(y,s) \approx Z_2(y;p)e^{ps}$, which is precisely what the evolution of the free mode should be. Second, the growing modes cannot be excited by the initial profile analytic at $y = 1$. If this were the case, then h could be expanded in a Taylor series around

$y = 1$, each term in the series raised to integer power corresponding to non-negative value of α, and so not growing at large s. Third, the above argument illustrates how only the Z_2 content of the initial data is relevant to the subsequent evolution of the wavepacket.

So it seems that the exponential growth of the wavepacket at late times must be already built into the initial data in the form of a power-law divergence of the initial wave profile, just as it is encoded in the pure free mode Z_2. But a power-law divergence of the perturbation near $v = 0$ causes curvature invariants to diverge at the junction, making the surface $v = 0$ a weak null singularity, which casts a shadow of doubt on the physicality of such growing modes. The question then arises whether it is possible to somehow eliminate the offending divergence, while still having a wavepacket grow to large enough values. The answer to this question is yes, and it is discussed next.

6.2.2 Evolution of a wavepacket initially localized at $v = 0$

What will happen if we cut off a diverging initial wave shape (6.28) below some small value of $y - 1$, say λ? To put it another way, will the power-law diverging wave localized near $y = 1$ backscatter and affect the evolution of the wavepacket at the large y? If not, then subtracting such a localized wavepacket from the perturbation modes discussed above will cut off the divergence at $y = 1$, while keeping the rest of the wavepacket evolution essentially unchanged.

We model such a localized wave by adding an exponential cutoff to the generic power law initial term (6.28)

$$h(\eta) \propto (1-\eta)^{\alpha} e^{(1-\eta)/\lambda}. \tag{6.34}$$

The exponential factor is chosen because it effectively suppresses h for values of $y-1 > \lambda$, yet still keeps the calculations simple. We are interested in the evolution of the wavepacket well outside the region of initial localization, but still for small enough y so that the approximation (6.26) holds, that is for $\lambda \ll y-1 \ll 1$ (which can always be arranged for small enough λ). In this region of interest we have

$$F(y,k) \approx \int_1^y \frac{1}{1-k}\left[\left(\frac{1-\eta}{1-y}\right)^{k-1} \mathcal{F}_2(y) - \mathcal{F}_1(y)\right](1-\eta)^{\alpha} e^{(1-\eta)/\lambda}\, d\eta. \tag{6.35}$$

The above integral can be evaluated by the change of variable

$$t = \frac{\eta - 1}{\lambda}, \tag{6.36}$$

which yields the following approximation for $F(y;k)$ in the region of interest:

$$F(y,k) \approx \int_0^\infty \frac{1}{1-k}\left[\left(\frac{-\lambda t}{1-y}\right)^{k-1}\mathcal{F}_2(y)-\mathcal{F}_1(y)\right](-\lambda t)^\alpha e^{-t}\,\lambda dt \qquad (6.37)$$

$$= -\frac{1}{k-1}\left[(-\lambda)^{k+\alpha}\Gamma(k+\alpha)(1-y)^{1-k}\mathcal{F}_2(y) - (-\lambda)^{1+\alpha}\Gamma(1+\alpha)\mathcal{F}_1(y)\right].$$

So, outside the region of initial localization of the wavepacket, but still for small y, the Laplace transform of the field perturbation is approximately given by

$$F(y,k) \approx -\frac{1}{k-1}\left[(-\lambda)^{k+\alpha}\Gamma(k+\alpha)Z_2(y) - (-\lambda)^{1+\alpha}\Gamma(1+\alpha)Z_1(y)\right]. \qquad (6.38)$$

The main contribution to the late-time behavior is coming from the poles of the gamma-function $\Gamma(k+\alpha)$ in the first term. Using the stationary phase approximation, it is possible to calculate this contribution exactly. The inverse Laplace transform of $F(y;k)$ can be reduced to the inverse Mellin transform of the gamma-function

$$f(y,s) \approx \frac{1}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty}\frac{Z_2(y;-\alpha)}{1+\alpha}\lambda^{k+\alpha}\Gamma(k+\alpha)e^{ks}dk \qquad (6.39)$$

$$= \frac{Z_2(y;-\alpha)}{1+\alpha}\frac{e^{-\alpha s}}{2\pi i}\int_{\kappa+\alpha-i\infty}^{\kappa+\alpha+i\infty}\lambda^k\Gamma(k)e^{ks}dk \qquad (6.40)$$

$$= \frac{Z_2(y;-\alpha)}{1+\alpha}\frac{e^{-\alpha s}}{2\pi i}\int_{\kappa+\alpha-i\infty}^{\kappa+\alpha+i\infty}\Gamma(k)\left[e^{-(s+\ln\lambda)}\right]^{-k}dk \qquad (6.41)$$

$$= \frac{Z_2(y;-\alpha)}{1+\alpha}e^{-\alpha s}\left[\mathcal{M}^{-1}\Gamma\right]\left(e^{-(s+\ln\lambda)}\right), \qquad (6.42)$$

which is a known integral; indeed, $[\mathcal{M}^{-1}\Gamma](x) = e^{-x}$. Therefore we obtain the following late-time approximation of the field perturbation outside the region of initial localization:

$$f(y,s) \approx \frac{\Theta(s+\ln\lambda)}{1+\alpha}Z_2(y;-\alpha)\,e^{-\alpha s}. \qquad (6.43)$$

This looks very similar to the earlier result (6.33); the only difference is the factor $\Theta(s+\ln\lambda)$, where Θ is defined by

$$\Theta(x) = \exp\left[-e^{-x}\right], \qquad (6.44)$$

and resembles a smoothed-out step function, rapidly changing its value from 0 to 1 as its argument becomes positive

$$\Theta(x) \approx \begin{cases} 1, & \mathrm{Re}\, x > 0 \\ 0, & \mathrm{Re}\, x < 0 \end{cases}, \tag{6.45}$$

with the width of the transition of order unity. This means that the perturbation outside the region of initial localization does not feel the effect of the field at $y - 1 < \lambda$ until a much later time, namely $s = -\ln\lambda$, when it suddenly spreads. To put it simply, the wavepacket initially localized at $v < \lambda$ does not backscatter until it hits the singularity at $u = 0$, and then goes out in a narrow band $-u < \lambda$.

The above result shows how we can cut off the free mode $Z_2(y;k)$ to avoid the curvature divergence, and yet have it grow sufficiently large. To quantify how large it can grow, consider the free mode $f = Z_2(y;k)e^{ks} \approx (1-y)^{1-k}e^{ks}$, with divergent scalar curvature $R \propto f'$, and cut it off at $y - 1 < \lambda$. The largest initial curvature value is of order $R \propto \lambda^{-k}$, while the initial energy of the pulse is $\lambda R \propto \lambda^{1-k}$, which can be made arbitrarily small. The perturbation mode will grow exponentially until the cutoff backscatters at $s = -\ln\lambda$, at which time its amplitude will be λ^{-k}, with proportionally large curvatures and energies. In other words, the initial large curvature seed localized in a small region spreads over the whole space in the course of the evolution, with the energy of the pulse growing correspondingly. Thus, the free perturbation modes considered here are physical and grow exponentially to very large amplitudes, certainly enough to leave the linear regime, and are therefore responsible for the evolution of the solution away from the Roberts one.

6.2.3 Generic initial conditions

We now turn our attention to the evolution of the wavepacket from generic initial conditions. It is reasonable to expect that completely generic initial conditions will have non-zero content of all perturbation modes present in the system, both growing and decaying, as given by equations (6.23) and (6.24),

$$f(y,s) = \int_{\kappa-i\infty}^{\kappa+i\infty} \left[W_2(k)Z_2(y;k) + W_1(k)Z_1(y;k)\right] e^{ks} dk. \tag{6.46}$$

However, decaying modes will disappear very quickly, so only the growing modes are relevant to late-time evolution. Assuming that the content of the free growing mode $Z_2(y;k)$ is given by the weight function $W(k)$, the generic wavepacket evolution is given by the sum of all such modes

$$f(y,s) = \int_{\Gamma} W(k)Z_2(y;k)e^{ks} dk, \tag{6.47}$$

where the infinite contour of integration Γ runs vertically in the regions F and $\bar{F}$ of the complex plane on Fig. 6, at Re $k = 1$. However, we note that the part of the contour between the endpoints $k_0^{\pm} = 1 \pm i\sqrt{2}$ of the regions F and $\bar{F}$ does not correspond to free growing modes, as $Z_2(y \to \infty) \approx (-y)^{-a}$ with Re $a < 0$ there, and so boundary conditions at infinity are not satisfied. Therefore, for the initial wavepacket bounded at infinity, the content of such modes is suppressed, so that $\int dk\,[W(k)\,(-y)^{-a}] \sim 1$ for large y. This leads to slower growth rates $\int dk\,[W(k)\,(-ye^s)^{-a} e^{(k+a)s}] \sim e^{(\kappa+a)s}$ at the later times. Hence, the piece of the contour Γ between the endpoints $k_0^{\pm}$ can be omitted from the integration without affecting the late-time evolution.

For completely generic initial conditions, we should expect W to be a smooth function of k in the free mode region F, not preferring any particular value of k. Therefore, using the stationary phase approximation, the main contribution to the late-time behavior of the above integral comes from the end points of the contour of integration,

$$f(y,s) \approx -W(k_0^+) Z_2(y;k_0^+)\,\frac{e^{k_0^+ s}}{s} - W(k_0^-) Z_2(y;k_0^-)\,\frac{e^{k_0^- s}}{s}. \tag{6.48}$$

Ignoring the overall weight factor, we find that the late-time evolution of the generic wavepacket is given by

$$f(y,s) \propto \mathrm{Re}\left[Z_2(y;k_0)\,\frac{e^{k_0 s}}{s}\right], \tag{6.49}$$

where k_0 can mean either k_0^+ or k_0^- — it does not matter as both give the same result. We emphasize that the single k_0-mode dominates the course of evolution of the generic wavepacket, and thus a certain universality is present in the way a generic perturbation departs from the Roberts solution.

6.3 Emergence of discrete self-similarity

As was shown above, the departure of the generic perturbation away from the Roberts solution is *universal* in a sense that the single mode $Z_2(y;k_0)e^{k_0 s}$ dominates the late-time evolution of the field. The complex growth exponent gives rise to an interesting physical effect: the perturbation developing on the scale-invariant background evolves to have a scale-dependent structure $e^s \cos(\mathrm{Im}\ k_0 s)$. The exponential growth of the amplitude of the perturbation will eventually be stopped by the non-linear effects, while the periodic dependence of the perturbation on the scale will most likely remain. The period of oscillation, obtained in the linear approximation, is

$$\Delta = \frac{2\pi}{\mathrm{Im}\ k_0}. \tag{6.50}$$

What does this periodic dependence of the solution on the scale mean physically? To answer this question, let us see how this symmetry is expressed in the Schwarzschild coordinates (r, t) often used in numerical calculations (see, for example, [31]). In these coordinates the Roberts metric (3.62) becomes diagonal

$$ds^2 = -\alpha\, dt^2 + \beta\, dr^2 + r^2\, d\Omega^2, \tag{6.51}$$

where metric coefficients and explicit expressions for coordinates were given in Section 3.6.4. For our purposes, it suffices to note that the coordinate x determines the ratio r/t, while the coordinate s sets an overall scale of both space and time coordinates via e^{-s} factor

$$-\frac{r}{t} = \exp\left[x + \frac{1}{2}\, e^{2x}\right], \quad r = \exp[x - s]. \tag{6.52}$$

One can see that taking a step Δ in the scale variable s is equivalent to scaling both spatial and time coordinates r and t down by a factor e^{Δ}. Therefore, the solution being periodic in scale coordinate s is equivalent to being invariant under rescaling of space and time coordinates r and t by a certain factor

$$f(x, s + \Delta) = f(x, s) \iff f(e^{-\Delta} r, e^{-\Delta} t) = f(r, t). \tag{6.53}$$

The latter is an expression of the symmetry observed in the numerical simulations of the massless scalar field collapse [31], and referred to as echoing, or discrete self-similarity, in the literature [31, 63].

Thus, our simple analytical model of the critical collapse of the massless scalar field illustrates how the continuous self-similarity of the Roberts solution is dynamically broken to discrete self-similarity by the growing perturbations, reproducing the essential feature of numerical critical solutions. The value of exponent for the endpoints of the spectrum of growing perturbation modes, $k_0 = 1 + i\sqrt{2}$, gives the period of discrete self-similarity as

$$\Delta = \sqrt{2}\pi = 4.44 \tag{6.54}$$

for linear perturbations of the Roberts solution, which is within 25% of the numerical value $\Delta = 3.44$ measured by Choptuik [31]. Given that the perturbative model considered here reproduces all the symmetries of the Choptuik's solution, and gives a good estimate for the period of echoing, it is instructive to compare the actual field profiles to the numerical calculations, which we will do next.

To meaningfully compare field profiles, we should cast the gauge-invariant quantities used throughout this thesis into the variables Choptuik uses, i.e. write the perturbation amplitudes in the gauge preserving diagonal form of the metric. The explicit expressions for field perturbations in this gauge turn out to be quite

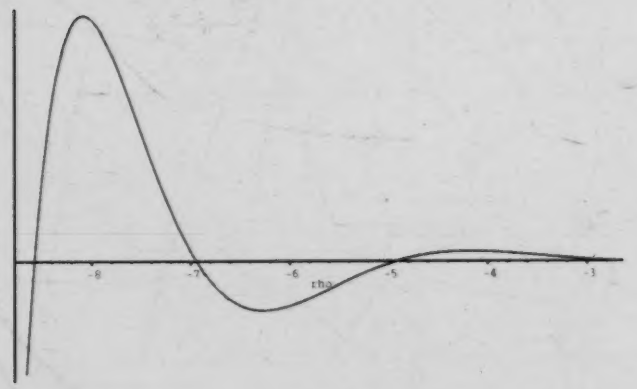


Figure 8: Profile for the field variable $X = \sqrt{2\pi}\sqrt{\frac{r^2}{a}}\frac{\partial\phi}{\partial r}$ on the slice $t = \text{const}$ for the dominant mode $f(y,s) = Z_2(y;k_0)\,e^{ks}/s$.

complicated; the simplest way to get them from gauge-invariant quantities is to find explicitly a gauge transformation

$$\xi^\mu = (A, B, 0, 0) \tag{6.55}$$

connecting the simple field gauge $K = k_{vv} = 0$ with the diagonal gauge, fixed by conditions $K = 0$ and $(2u - v)^2 k_{vv} = u^2 k_{uu}$. The effects of the gauge transformation on the perturbation amplitudes were given in Section 4.1. Imposing the condition $K = 0$, one finds that B must be related to A by

$$B = \frac{2u - v}{u} A. \tag{6.56}$$

A is then found by imposing the other condition fixing the diagonal gauge, which leads to the following equation

$$(2u - v)^2 A_{,v} - u(2u - v)A_{,u} - vA = 2u \int f\, dv. \tag{6.57}$$

Rewriting A in scaling coordinates,

$$A(y,s) = \mathcal{A}(y) e^{(k-1)s}, \tag{6.58}$$

transforms the above equation into the ordinary differential equation

$$(1 + y)\dot{\mathcal{A}} + [1 - k/2\,(1 + y)]\mathcal{A} = -\int F\, dy, \tag{6.59}$$

which can be solved to give

$$\mathcal{A}(y) = -\frac{e^{\frac{k}{2}y}}{1+y} \int_1^y d\xi\, e^{-\frac{k}{2}\xi} \int_1^\xi F(\zeta)\, d\zeta. \tag{6.60}$$

Once the connecting gauge transformation is known, we can obtain the perturbation amplitudes in the Schwarzschild diagonal gauge. In particular, the scalar field perturbation is given by

$$\varphi(y;k) = -F(y;k) + \mathcal{A}(y;k). \tag{6.61}$$

The end result of calculation for the field variable $X = \sqrt{2\pi}\sqrt{\frac{r^2}{a}}\frac{\partial\phi}{\partial r}$ from the perturbation modes is presented in Fig. 8. Comparing this plot to Fig. 2 in Choptuik's original paper [31], we see that they share one common feature, which is the oscillatory nature of the field solution; however, the shape of the field profiles is quite different. This discrepancy is not surprising, perhaps, since perturbation methods in critical phenomena are usually viable for calculating critical exponents, but not the field configurations themselves.

We note that the above discussion was based on linear perturbation analysis. Growing perturbation modes will eventually become sufficiently big, and will take the solution into the non-linear regime. It is likely the discretely self-similar structure would be preserved, but the exact value of the period and the shape of the field profiles could be modified. The non-linear regime requires more investigation.

Chapter 7

Discussion

7.1 Our results

We have searched for and found continuously self-similar spherically symmetric solutions of minimally coupled scalar field collapse in n-dimensional spacetimes. For spacetime dimensions higher than three they form a one-parameter family and display critical behavior much like the Roberts solution. The qualitative picture of field evolution is easy to visualize in analogy with a particle traveling in a potential of inverted U shape. This analogy was used in [5, 6] to consider black hole formation in subcritical collapse by quantum tunneling. Critical solutions are in general simpler than other members of the family due to the potential factoring.

We also used the equivalence of scalar field couplings to generalize solutions of the minimally coupled scalar field to a much wider class of couplings. For often-used cases of conformal coupling and dilaton gravity the results are remarkably simple. Some results of [73], applied for a single scalar field only, become trivial in view of this coupling equivalence.

We carried out a linear perturbation analysis of the Roberts solution in a general gauge-invariant formalism to investigate the critical behavior in the self-similar gravitational collapse of a massless scalar field. An exact analysis of the perturbation eigenvalue problem reveals that there are no growing non-spherical perturbation modes. However, there are growing spherical perturbation modes. Their spectrum is continuous and occupies a big chunk of the complex plane,

$$\frac{1}{2} < \mathrm{Re}\, k < 1, \quad |\mathrm{Im}\, k| > \sqrt{\frac{\mathrm{Re}\, k\,(2 - \mathrm{Re}\, k)}{1 - (2\mathrm{Re}\, k)^{-1}}}. \tag{7.1}$$

In view of these findings, the following picture of the dynamics of scalar field evolution near self-similarity emerges: As we evolve generic initial data which is sufficiently close to the critical Roberts solution, non-spherical modes decay

and the solution approaches the spherically symmetric one. Asymmetry of the initial data does not play a role in the collapse. The growing spherical modes, on the other hand, drive the solution farther away from the continuously self-similar one. In this sense, the critical Roberts solution is an intermediate attractor for non-spherical initial data.

A continuous perturbation spectrum could possibly suggest non-universality of the critical behavior for different ingoing wavepackets. To investigate this question, we studied spherically symmetric perturbations of the Roberts solution with the intent to understand how nearby solutions depart from the Roberts one in the course of the field evolution. We analyzed the behavior of incoming and outgoing wavepackets, and we focused our attention on the incoming one as the physically relevant one for the question posed. With the aid of the Green's function formulation, we were able to solve the perturbation problem completely in closed form, as well as obtain simple approximations for the late-time evolution of the field in several important regimes. We discovered that discretely self-similar structure forms on the continuously self-similar background in the collapse of a minimally-coupled massless scalar field, and we calculated the period of this structure to be

$$\Delta = \sqrt{2}\pi = 4.44. \tag{7.2}$$

7.2 Further studies

An interesting question, which is not answered by perturbative calculations, is the further fate of the scalar field evolution as it gets away from the Roberts solution. The emerging discretely self-similar structure, and the universal way in which the generic perturbation departs from the Roberts solution, offer support for the conjecture that the Roberts solution is "close" to the Choptuik one in the phase space of all massless scalar field configurations (in a sense of being in the basin of attraction of the latter), and will evolve towards it when perturbed. It seems highly unlikely, however, that the critical mode responsible for the decay of the Choptuik solution will be completely absent in the initial data originating near the Roberts solution, as this usually requires fine-tuning of the parameters. So, although at first the field configuration near the Roberts solution might seem to evolve towards the Choptuik solution, after a while the critical mode will kick in and drive the field to either dispersal or black hole formation.

This picture is in line with the Choptuik solution being an *intermediate* attractor. One can think of the field evolution in a phase space as a stone rolling down a mountain range, with local attractors being mountain passes. As a stone rolls down to the lower elevations, it would follow valleys and gorges to get from a

higher pass to a lower pass, until finally it reaches the bottom. In this analogy, we can speculate that the Choptuik solution lies downstream from the Roberts solution, and the field evolution of nearly continuously self-similar data would follow the path from an initial configuration near the Roberts solution to one near the Choptuik solution and then to a global attractor (black hole or flat spacetime) in the phase space. Unfortunately, linear perturbation methods are not sufficient to provide a proof of the proposed scenario, and they fail to give the answer as to what would the eventual fate of the evolution be (whether black hole or flat spacetime end-state will be selected), and how fast would the field get there.

To completely answer these questions, one would need to employ some sort of non-linear calculation, or perform numerical simulations of the evolution. In particular, it would be interesting to evolve the perturbed Roberts spacetime numerically and look for the Choptuik spacetime as the possible intermediate attractor. Nevertheless, it may happen that some information about Choptuik's solution can be gained from the linear perturbation analysis of the Roberts solution. The appeal of this method lies in the fact that such analysis could be carried out analytically, while Choptuik's solution is still unknown in closed form. Similarly, one can try to study properties of other analytically-unknown critical solutions in different matter models based on "nearby" solutions with higher symmetry and simpler form. This proposal was recently implemented in Ref. [69]. One might also hope to obtain acceptable analytical approximations to critical solutions, Choptuik's in particular, by going to higher order perturbation theory in the region near the singularity.

Appendices

Appendix A

Properties of the Hypergeometric Equation

In this thesis, we argued that the linear perturbation analysis of the Roberts solution can be reduced to the study of solutions of the hypergeometric equation with certain parameters. The hypergeometric equation has been extensively studied: for a complete description of its main properties see, for example, [8]. In this appendix we collect the facts about the hypergeometric equations that are of immediate use to us, mainly to establish notation.

The hypergeometric equation is a second order linear ordinary differential equation,

$$y(1-y)\ddot{Z} + [c-(a+b+1)y]\dot{Z} - abZ = 0, \tag{A.1}$$

with parameters a, b, and c being arbitrary complex numbers. It has three singular points at $y = 0, 1, \infty$. Its general solution is a linear combination of any two different solutions from the set

$$\begin{aligned}
Z_1 &= \mathcal{F}(a,b;a+b+1-c;1-y), \\
Z_2 &= (1-y)^{c-a-b}\mathcal{F}(c-a,c-b;c+1-a-b;1-y), \\
Z_3 &= (-y)^{-a}\mathcal{F}(a,a+1-c;a+1-b;y^{-1}), \\
Z_4 &= (-y)^{-b}\mathcal{F}(b+1-c,b;b+1-a;y^{-1}), \\
Z_5 &= \mathcal{F}(a,b;c;y), \\
Z_6 &= y^{1-c}\mathcal{F}(a+1-c,b+1-c;2-c;y),
\end{aligned} \tag{A.2}$$

where $\mathcal{F}(a,b;c;y)$ is the hypergeometric function, defined by the power series

$$\mathcal{F}(a,b;c;y) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{y^n}{n!}, \tag{A.3}$$

and we used shorthand notation $(a)_n = \Gamma(a+n)/\Gamma(a)$. The hypergeometric series is regular at $y = 0$, its value there is $\mathcal{F}(a, b; c; 0) = 1$, and it is absolutely convergent for $|y| < 1$. Considering the hypergeometric series as a function of its parameters, one can show that $\mathcal{F}(a, b; c; y_0)/\Gamma(c)$ is an entire analytical function of a, b, and c, provided that $|y_0| < 1$.

The solutions $Z_1, \ldots, Z_6$ are based around different singular points of the hypergeometric equation, with asymptotics given by

$$\begin{aligned} Z_1 = 1,\ Z_2 = (1-y)^{c-a-b} \quad &\text{near} \quad y = 1, \\ Z_3 = (-y)^{-a},\ Z_4 = (-y)^{-b} \quad &\text{near} \quad y = \infty, \\ Z_5 = 1,\ Z_6 = y^{1-c} \quad &\text{near} \quad y = 0. \end{aligned} \tag{A.4}$$

Any three of the functions $Z_1, \ldots, Z_6$ are linearly dependent with constant coefficients. In particular,

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} c_{13} & c_{14} \\ c_{23} & c_{24} \end{bmatrix} \begin{bmatrix} Z_3 \\ Z_4 \end{bmatrix}, \tag{A.5}$$

where the coefficient matrix is given by

$$\begin{bmatrix} c_{13} & c_{14} \\ c_{23} & c_{24} \end{bmatrix} = \begin{bmatrix} \frac{\Gamma(a+b+1-c)\Gamma(b-a)}{\Gamma(b+1-c)\Gamma(b)} e^{-i\pi a} & \frac{\Gamma(a+b+1-c)\Gamma(a-b)}{\Gamma(a+1-c)\Gamma(a)} e^{-i\pi b} \\ \frac{\Gamma(c+1-a-b)\Gamma(b-a)}{\Gamma(1-a)\Gamma(c-a)} e^{-i\pi(c-b)} & \frac{\Gamma(c+1-a-b)\Gamma(a-b)}{\Gamma(1-b)\Gamma(c-b)} e^{-i\pi(c-a)} \end{bmatrix}. \tag{A.6}$$

These relationships, known as Kummer series, are true for all values of the parameters for which the gamma-function terms in the numerators are finite, and all values of y for which corresponding series converge, with $\mathrm{Im}\, y > 0$. If $\mathrm{Im}\, y < 0$, signs of arguments in the exponential multipliers should be inverted. We shall not give the remaining similar relationships here.

Bibliography

[1] A. M. Abrahams and C. R. Evans, *Critical behavior and scaling in vacuum axisymmetric gravitational collapse*, Phys. Rev. Lett. **70**, 2980 (1993).

[2] A. M. Abrahams and C. R. Evans, *Critical phenomena and relativistic gravitational collapse*, Gen. Rel. Grav. **26**, 379 (1994).

[3] M. Alcubierre, G. Allen, B. Brügmann, and G. Lanfermann, *Gravitational collapse of gravitational waves in 3D numerical relativity*, Phys. Rev. D **61**, 041501 (2000), gr-qc/9904013.

[4] S. Ayal and T. Piran, *Spherical collapse of a mass-less scalar field with semi-classical corrections*, Phys. Rev. D **56**, 4768 (1997), gr-qc/9704027.

[5] D. Bak, S.-P. Kim, S.-K. Kim, K.-S. Soh, and J.-H. Yee, *Quantum creation of black hole by tunneling in scalar field collapse*, Phys. Rev. D **60**, 064005 (1999), gr-qc/9901028.

[6] D. Bak, S.-P. Kim, S.-K. Kim, K.-S. Soh, and J.-H. Yee, *Wave functions for quantum black hole formation in scalar field collapse*, Phys. Rev. D **61**, 044005 (2000), gr-qc/9907032.

[7] C. Barrabes and W. Israel, *Thin shells in general relativity and cosmology: The Lightlike limit*, Phys. Rev. D **43**, 1129 (1991).

[8] H. Bateman and A. Erdélyi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953).

[9] J. D. Bekenstein, *Exact solutions of Einstein – conformal scalar equations*, Ann. Phys. **82**, 535 (1974).

[10] D. Birmingham and S. Sen, *Gott time machines, BTZ black hole formation, and Choptuik scaling*, Phys. Rev. Lett. **84**, 1074 (2000), hep-th/9908150.

[11] P. Bizoń, *How to make a tiny black hole?*, Acta Cosmologica **22**, 81 (1996), gr-qc/9606060.

[12] P. Bizoń and T. Chmaj, *Critical collapse of skyrmions*, Phys. Rev. D **58**, 041501 (1998), gr-qc/9801012.

[13] P. Bizoń and T. Chmaj, *First order phase transitions in gravitational collapse*, Acta Phys. Polon. **B29**, 1071 (1998), gr-qc/9802002.

[14] P. Bizoń, T. Chmaj, and Z. Tabor, *On equivalence of critical collapse of non-Abelian fields*, Phys. Rev. D **59**, 104003 (1999), gr-qc/9901039.

[15] W. B. Bonnor and M. S. Piper, *Implosion of quadrupole gravitational waves* (1996), gr-qc/9610015.

[16] S. Bose, L. Parker, and Y. Peleg, *Predictability and semiclassical approximation at the onset of black hole formation*, Phys. Rev. D **54**, 7490 (1996), hep-th/9606152.

[17] P. R. Brady, *Does scalar field collapse produce "zero mass" black holes?* (1994), gr-qc/9402023.

[18] P. R. Brady, *Selfsimilar scalar field collapse: Naked singularities and critical behavior*, Phys. Rev. D **51**, 4168 (1995), gr-qc/9409035.

[19] P. R. Brady, C. M. Chambers, and S. M. C. V. Gonçalves, *Phases of massive scalar field collapse*, Phys. Rev. D **56**, 6057 (1997), gr-qc/9709014.

[20] P. R. Brady and A. C. Ottewill, *Quantum corrections to critical phenomena in gravitational collapse*, Phys. Rev. D **58**, 024006 (1998), gr-qc/9804058.

[21] P. R. Brady and M. J. Cai, *Critical phenomena in gravitational collapse* (1998), gr-qc/9812071.

[22] L. M. Burko, *Comment on the Roberts solution for the spherically-symmetric Einstein-scalar field equations*, Gen. Rel. Grav. **29**, 259 (1997), gr-qc/9608061.

[23] L. M. Burko, *The singularity in supercritical collapse of a spherical scalar field*, Phys. Rev. D **58**, 084013 (1998), gr-qc/9803059.

[24] B. J. Carr and A. A. Coley, *Self-Similarity in General Relativity* (1998), gr-qc/9806048.

[25] B. J. Carr, A. A. Coley, M. Goliath, U. S. Nilsson, and C. Uggla, *Critical phenomena and a new class of self-similar spherically symmetric perfect-fluid solutions*, Phys. Rev. D **61**, 081502 (2000), gr-qc/9901031.

[26] B. J. Carr and A. A. Coley, *A complete classification of spherically symmetric perfect fluid similarity solutions*, Phys. Rev. D **62**, 044023 (2000), gr-qc/9901050.

[27] B. J. Carr, A. A. Coley, M. Goliath, U. S. Nilsson, and C. Uggla, *Physical interpretation of self-similar spherically symmetric perfect-fluid models — combining the comoving and homothetic approach* (1999), gr-qc/9902070.

[28] C. M. Chambers, P. R. Brady, and S. M. C. V. Gonçalves, *A critical look at massive scalar field collapse* (1997), gr-qc/9710014.

[29] T. Chiba and M. Siino, *Disappearance of critical behavior in semiclassical general relativity* (1996), KUNS-1384.

[30] T. Chiba and J. Soda, *Critical behavior in the Brans-Dicke theory of gravitation*, Prog. Theor. Phys. **96**, 567 (1996), gr-qc/9603056.

[31] M. W. Choptuik, *Universality and scaling in gravitational collapse of a massless scalar field*, Phys. Rev. Lett. **70**, 9 (1993).

[32] M. W. Choptuik, T. Chmaj, and P. Bizoń, *Critical Behaviour in Gravitational Collapse of a Yang-Mills Field*, Phys. Rev. Lett. **77**, 424 (1996), gr-qc/9603051.

[33] M. W. Choptuik, E. W. Hirschmann, and S. L. Liebling, *Instability of an "approximate black hole"*, Phys. Rev. D **55**, 6014 (1997), gr-qc/9701011.

[34] M. W. Choptuik, *The (unstable) threshold of black hole formation* (1998), gr-qc/9803075.

[35] M. W. Choptuik, E. W. Hirschmann, and R. L. Marsa, *New critical behavior in Einstein-Yang-Mills collapse*, Phys. Rev. D **60**, 124011 (1999), gr-qc/9903081.

[36] D. M. Eardley, E. W. Hirschmann, and J. H. Horne, *S-duality at the black hole threshold in gravitational collapse*, Phys. Rev. D **52**, 5397 (1995), gr-qc/9505041.

[37] C. R. Evans and J. S. Coleman, *Observation of critical phenomena and self-similarity in the gravitational collapse of radiation fluid*, Phys. Rev. Lett. **72**, 1782 (1994), gr-qc/9402041.

[38] L. H. Ford and L. Parker, *Creation of particles by singularities in asymptotically flat space-times*, Phys. Rev. D **17**, 1485 (1978).

[39] A. V. Frolov, *Perturbations and critical behavior in the self-similar gravitational collapse of a massless scalar field*, Phys. Rev. D **56**, 6433 (1997), gr-qc/9704040.

[40] A. V. Frolov, *Self-similar collapse of scalar field in higher dimensions*, Class. Quant. Grav. **16**, 407 (1999), gr-qc/9806112.

[41] A. V. Frolov, *Critical collapse beyond spherical symmetry: General perturbations of the Roberts solution*, Phys. Rev. D **59**, 104011 (1999), gr-qc/9811001.

[42] A. V. Frolov, *Continuous self-similarity breaking in critical collapse*, Phys. Rev. D **61**, 084006 (2000), gr-qc/9908046.

[43] V. P. Frolov, A. L. Larsen, and M. Christensen, *Domain wall interacting with a black hole: A new example of critical phenomena*, Phys. Rev. D **59**, 125008 (1999), hep-th/9811148.

[44] D. Garfinkle, *Choptuik scaling in null coordinates*, Phys. Rev. D **51**, 5558 (1995), gr-qc/9412008.

[45] D. Garfinkle, *Choptuik scaling and the scale invariance of Einstein's equation*, Phys. Rev. D **56**, 3169 (1997), gr-qc/9612015.

[46] D. Garfinkle and G. C. Duncan, *Scaling of curvature in sub-critical gravitational collapse*, Phys. Rev. D **58**, 064024 (1998), gr-qc/9802061.

[47] D. Garfinkle and K. Meyer, *Scale invariance and critical gravitational collapse*, Phys. Rev. D **59**, 064003 (1999), gr-qc/9806052.

[48] D. Garfinkle, C. Gundlach, and J. M. Martin-Garcia, *Angular momentum near the black hole threshold in scalar field collapse*, Phys. Rev. D **59**, 104012 (1999), gr-qc/9811004.

[49] D. Garfinkle and C. Gundlach, *Symmetry-seeking spacetime coordinates*, Class. Quant. Grav. **16**, 4111 (1999), gr-qc/9908016.

[50] D. Garfinkle, C. Cutler, and G. C. Duncan, *Choptuik scaling in six dimensions*, Phys. Rev. D **60**, 104007 (1999), gr-qc/9908044.

[51] U. H. Gerlach and U. K. Sengupta, *Gauge-invariant perturbations on most general spherically symmetric space-times*, Phys. Rev. D **19**, 2268 (1979).

[52] U. H. Gerlach and U. K. Sengupta, *Gauge-invariant coupled gravitational, acoustical, and electromagnetic modes on most general spherical space-times*, Phys. Rev. D **22**, 1300 (1980).

[53] M. Goliath, U. S. Nilsson, and C. Uggla, *Spatially self-similar spherically symmetric perfect-fluid models*, Class. Quant. Grav. **15**, 167 (1998), gr-qc/9811064.

[54] M. Goliath, U. S. Nilsson, and C. Uggla, *Timelike self-similar spherically symmetric perfect-fluid models*, Class. Quant. Grav. **15**, 2841 (1998), gr-qc/9811065.

[55] S. M. C. V. Gonçalves and I. G. Moss, *Black hole formation from massive scalar fields*, Class. Quant. Grav. **14**, 2607 (1997), gr-qc/9702059.

[56] A. M. Green and A. R. Liddle, *Critical collapse and the primordial black hole initial mass function*, Phys. Rev. D **60**, 063509 (1999), astro-ph/9901268.

[57] C. Gundlach, *The Choptuik space-time as an eigenvalue problem*, Phys. Rev. Lett. **75**, 3214 (1995), gr-qc/9507054.

[58] C. Gundlach, *Understanding critical collapse of a scalar field*, Phys. Rev. D **55**, 695 (1997), gr-qc/9604019.

[59] C. Gundlach and J. M. Martín-García, *Charge scaling and universality in critical collapse*, Phys. Rev. D **54**, 7353 (1996), gr-qc/9606072.

[60] C. Gundlach, *Echoing and scaling in Einstein-Yang-Mills critical collapse*, Phys. Rev. D **55**, 6002 (1997), gr-qc/9610069.

[61] C. Gundlach, *Nonspherical perturbations of critical collapse and cosmic censorship*, Phys. Rev. D **57**, 7075 (1998), gr-qc/9710066.

[62] C. Gundlach, *Angular momentum at the black hole threshold*, Phys. Rev. D **57**, 7080 (1998), gr-qc/9711079.

[63] C. Gundlach, *Critical phenomena in gravitational collapse*, Adv. Theor. Math. Phys. **2**, 1 (1998), gr-qc/9712084.

[64] C. Gundlach, *Critical gravitational collapse of a perfect fluid with $p = k\rho$: Nonspherical perturbations* (1999), gr-qc/9906124.

[65] R. S. Hamadé and J. M. Stewart, *The Spherically symmetric collapse of a massless scalar field*, Class. Quant. Grav. **13**, 497 (1996), gr-qc/9506044.

[66] R. S. Hamadé, J. H. Horne, and J. M. Stewart, *Continuous Self-Similarity and S-Duality*, Class. Quant. Grav. **13**, 2241 (1996), gr-qc/9511024.

[67] T. Hara, T. Koike, and S. Adachi, *Renormalization group and critical behaviour in gravitational collapse* (1996), gr-qc/9607010.

[68] T. Harada, *Final fate of the spherically symmetric collapse of a perfect fluid*, Phys. Rev. D **58**, 104015 (1998), gr-qc/9807038.

[69] S. A. Hayward, *An extreme critical space-time: echoing and black-hole perturbations* (2000), gr-qc/0004038.

[70] E. W. Hirschmann and D. M. Eardley, *Universal scaling and echoing in gravitational collapse of a complex scalar field*, Phys. Rev. D **51**, 4198 (1995), gr-qc/9412066.

[71] E. W. Hirschmann and D. M. Eardley, *Critical exponents and stability at the black hole threshold for a complex scalar field*, Phys. Rev. D **52**, 5850 (1995), gr-qc/9506078.

[72] E. W. Hirschmann and D. M. Eardley, *Criticality and Bifurcation in the Gravitational Collapse of a Self-Coupled Scalar Field*, Phys. Rev. D **56**, 4696 (1997), gr-qc/9511052.

[73] E. W. Hirschmann and A. Wang, *Spherical self-similar solutions in Einstein-multi-scalar gravity*, Phys. Lett. A **249**, 383 (1998), gr-qc/9802065.

[74] S. Hod and T. Piran, *Fine-structure of Choptuik's mass-scaling relation*, Phys. Rev. D **55**, 440 (1997), gr-qc/9606087.

[75] S. Hod and T. Piran, *Critical behaviour and universality in gravitational collapse of a charged scalar field*, Phys. Rev. D **55**, 3485 (1997), gr-qc/9606093.

[76] V. Husain, E. A. Martinez, and D. Núñez, *Exact solution for scalar field collapse*, Phys. Rev. D **50**, 3783 (1994), gr-qc/9402021.

[77] V. Husain, E. A. Martinez, and D. Núñez, *On critical behavior in gravitational collapse*, Class. Quant. Grav. **13**, 1183 (1996), gr-qc/9505024.

[78] Y. Kiem, *Phase transition in spherically symmetric gravitational collapse of a massless scalar field* (1994), hep-th/9407100.

[79] Y. Kiem and D. Park, *Static and dynamic analysis of a massless scalar field coupled with a class of gravity theories* (1995), hep-th/9504021.

[80] T. Koike and T. Mishima, *An analytic model with critical behavior in black hole formation*, Phys. Rev. D **51**, 4045 (1995), gr-qc/9409045.

[81] T. Koike, T. Hara, and S. Adachi, *Critical behavior in gravitational collapse of radiation fluid: A Renormalization group (linear perturbation) analysis*, Phys. Rev. Lett. **74**, 5170 (1995), gr-qc/9503007.

[82] T. Koike, T. Hara, and S. Adachi, *Critical behavior in gravitational collapse of a perfect fluid*, Phys. Rev. D **59**, 104008 (1999).

[83] G. D. Kribs, A. K. Leibovich, and I. Z. Rothstein, *Bounds from Primordial Black Holes with a Near Critical Collapse Initial Mass Function*, Phys. Rev. D **60**, 103510 (1999), astro-ph/9904021.

[84] S. L. Liebling and M. W. Choptuik, *Black hole criticality in the Brans-Dicke model*, Phys. Rev. Lett. **77**, 1424 (1996), gr-qc/9606057.

[85] S. L. Liebling, *Multiply unstable black hole critical solutions*, Phys. Rev. D **58**, 084015 (1998), gr-qc/9805043.

[86] S. L. Liebling, *Critical phenomena inside global monopoles*, Phys. Rev. D **60**, 061502 (1999), gr-qc/9904077.

[87] C. O. Lousto, *Effective two-dimensional description from critical phenomena in black holes*, Gen. Rel. Grav. **27**, 121 (1995).

[88] D. Maison, *Non-universality of critical behavior in spherically symmetric gravitational collapse*, Phys. Lett. B **366**, 82 (1996), gr-qc/9504008.

[89] T. Maki and K. Shiraishi, *Exact solutions for gravitational collapse with a dilaton field in arbitrary dimensions*, Class. Quant. Grav. **12**, 159 (1995).

[90] R. L. Marsa and M. W. Choptuik, *Black hole - scalar field interactions in spherical symmetry*, Phys. Rev. D **54**, 4929 (1996), gr-qc/9607034.

[91] J. M. Martín-García and C. Gundlach, *All nonspherical perturbations of the Choptuik spacetime decay*, Phys. Rev. D **59**, 064031 (1999), gr-qc/9809059.

[92] J. P. Muniain and D. D. Piriz, *Critical behavior of dimensionally continued black holes*, Phys. Rev. D **53**, 816 (1996), gr-qc/9502029.

[93] D. W. Neilsen and M. W. Choptuik, *Critical phenomena in perfect fluids*, Class. Quant. Grav. **17**, 761 (2000), gr-qc/9812053.

[94] D. W. Neilsen and M. W. Choptuik, *Ultrarelativistic fluid dynamics*, Class. Quant. Grav. **17**, 733 (2000), gr-qc/9904052.

[95] J. C. Niemeyer and K. Jedamzik, *Near-Critical Gravitational Collapse and the Initial Mass Function of Primordial Black Holes*, Phys. Rev. Lett. **80**, 5481 (1998), astro-ph/9709072.

[96] U. Nilsson and C. Uggla, *Spatially self-similar locally rotationally symmetric perfect fluid models*, Class. Quant. Grav. **13**, 1601 (1996), gr-qc/9511064.

[97] D. Nunez, H. Quevedo, and M. Salgado, *Dynamics of a spherically symmetric scalar shell*, Phys. Rev. D **58**, 083506 (1998).

[98] H. P. de Oliveira and E. S. Cheb-Terrab, *Self-Similar Collapse of Conformally Coupled Scalar Fields*, Class. Quant. Grav. **13**, 425 (1996), gr-qc/9512010.

[99] H. P. de Oliveira, *Self-Similar Collapse in Brans-Dicke Theory and Critical Behavior* (1996), gr-qc/9605008.

[100] A. Ori and T. Piran, *Naked singularities in self-similar spherical gravitational collapse*, Phys. Rev. Lett. **59**, 2137 (1987).

[101] A. Ori and T. Piran, *Naked singularities and other features of self-similar general relativistic gravitational collapse*, Phys. Rev. D **42**, 1068 (1990).

[102] Y. Oshiro, K. Nakamura, and A. Tomimatsu, *Critical behavior of black hole formation in a scalar wave collapse*, Prog. Theor. Phys. **91**, 1265 (1994), gr-qc/9402017.

[103] Y. Oshiro, K. Nakamura, and A. Tomimatsu, *Critical behavior near the singularity in a scalar field collapse*, Phys. Rev. D **54**, 1540 (1996), dPNU-96-25.

[104] Y. Peleg and A. R. Steif, *Phase transition for gravitationally collapsing dust shells in 2+1 dimensions*, Phys. Rev. D **51**, 3992 (1995), gr-qc/9412023.

[105] Y. Peleg, S. Bose, and L. Parker, *Choptuik scaling and quantum effects in 2D dilaton gravity*, Phys. Rev. D **55**, 4525 (1997), gr-qc/9608040.

[106] U.-L. Pen, *A General class of self-similar self-gravitating fluids* (1994), astro-ph/9402039.

[107] R. H. Price and J. Pullin, *Analytic approximations to the spacetime of a critical gravitational collapse*, Phys. Rev. D **54**, 3792 (1996), gr-qc/9601009.

[108] J. Pullin, *Is there a connection between no hair behavior and universality in gravitational collapse?*, Phys. Lett. A **204**, 7 (1995), gr-qc/9409044.

[109] G. Rein, A. D. Rendall, and J. Schaeffer, *Critical collapse of collisionless matter: A numerical investigation*, Phys. Rev. D **58**, 044007 (1998), gr-qc/9804040.

[110] M. D. Roberts, *Scalar field counterexamples to the cosmic censorship hypothesis*, Gen. Rel. Grav. **21**, 907 (1989).

[111] T. P. Singh, *Gravitational Collapse and Cosmic Censorship* (1996), gr-qc/9606016.

[112] T. P. Singh, *Gravitational collapse, black holes and naked singularities* (1998), gr-qc/9805066.

[113] J. Soda and K. Hirata, *Higher dimensional self-similar spherical symmetric scalar field collapse and critical phenomena in black hole formation*, Phys. Lett. B **387**, 271 (1996), gr-qc/9605023.

[114] A. Strominger and L. Thorlacius, *Universality and scaling at the onset of quantum black hole formation*, Phys. Rev. Lett. **72**, 1584 (1994), hep-th/9312017.

[115] J. Traschen, *Discrete self-similiarity and critical point behavior in fluctuations about extremal black holes*, Phys. Rev. D **50**, 7144 (1994), gr-qc/9403016.

[116] J. F. Villas da Rocha, A. Wang, and N. O. Santos, *Gravitational Collapse of Perfect Fluid*, Phys. Lett. A **255**, 213 (1999), gr-qc/9811057.

[117] J. F. Villas da Rocha and A. Wang, *Gravitational Collapse of Perfect Fluid in N-Dimensional Spherically Symmetric Spacetimes* (1999), gr-qc/9910109.

[118] R. M. Wald, *Gravitational collapse and cosmic censorship* (1997), gr-qc/9710068.

[119] A. Wang and H. P. de Oliveira, *Critical phenomena of collapsing massless scalar wave packets*, Phys. Rev. D **56**, 753 (1997), gr-qc/9608063.

[120] A. Wang, J. F. Villas da Rocha, and N. O. Santos, *Gravitational collapse of massless scalar field and radiation fluid*, Phys. Rev. D **56**, 7692 (1997), gr-qc/9702057.

[121] A. Wang, *Critical collapse in tensor multiscalar and nonlinear gravity theories: A universal class* (1999), gr-qc/9901044.

[122] J. Yokoyama, *Cosmological constraints on primordial black holes produced in the near-critical gravitational collapse*, Phys. Rev. D **58**, 107502 (1998), gr-qc/9804041.